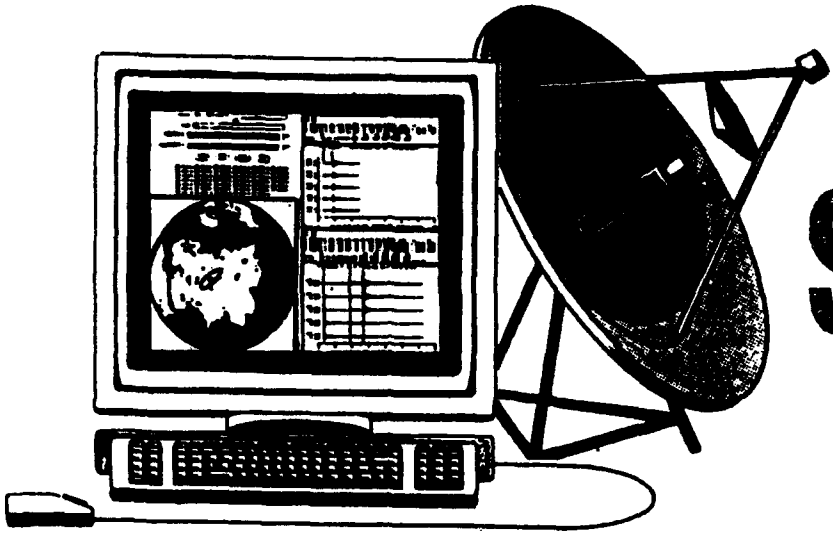
December 1993 SAIC-93/1219

Initial Wave-Type Identification with Neural Networks and its Contribution to Automated Processing in IMS Version 3.0



SPECIAL TECHNICAL REPORT
10 December 1993

*Thomas J. Sereno, Jr. and Gagan B. Patnaik**

Geophysical Systems Operations
Science Applications International Corporation

* Now at:
Advanced Geocomputing Technologies
P.O. Box 927477
San Diego, California 92192-7477

Sponsored by:
ADVANCED RESEARCH PROJECTS AGENCY
Nuclear Monitoring Research Office
ARPA Order Number 6266, Program Code No. 62714E
Issued by: ARPA/CMO
Contract No. MDA972-92-C0026

Principal Investigator:
Dr. Thomas C. Bache
(619) 458-2531
10260 Campus Point Drive
San Diego, California 92121

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	10 December 1993	Technical Report Nov. 1992-Dec. 1993

4. TITLE AND SUBTITLE
Initial Wave-type Identification with Neural Networks and it s Contribution to Automated Processing in IMS Version 3.0

5. FUNDING NUMBERS
MDA972-92-C-0026

6. AUTHOR(S)
Thomas J. Sereno, Jr. and Gagan B. Patnaik

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Science Applications International Corporation (SAIC)
10260 Campus Pt. Drive
San Diego, CA 92121

8. PERFORMING ORGANIZATION REPORT NUMBER
SAIC-93/1219

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Advanced Research Projects Agency (ARPA)
3701 N. Fairfax Drive, #717
Arlington, VA 22203-1714

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, Distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

(See reverse)

14. SUBJECT TERMS
Neural Networks, Seismology, Wave-type Identification

15. NUMBER OF PAGES
37

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Same as report

Abstract

This report describes a new 4-class neural network for automated identification of initial wave type (Teleseism, Regional *P*, Regional *S*, or Noise) for data recorded by 3-component stations or arrays. This is an extension of the 2-class (*P* or *S*) neural network that we developed for 3-component stations [*Patnaik and Sereno*, 1991]. The input data are dominant period, polarization attributes, contextual information (e.g., measurements related to a group of arrivals), a spectral representation of the horizontal-to-vertical power ratio, and the slowness determined by *f-k* analysis for array stations. We used a three-staged approach, and each stage consists of a 2-class neural network. The first stage separates signal from noise. The signals are passed to the second stage which separates regional *S* phases from regional *P* phases and teleseisms. The regional *P* phases and teleseisms are passed to the final stage which separates them into two distinct classes. A three-layer backpropagation neural network is used at each stage. Neural networks were trained for six 3-component IRIS/IDA stations in the CIS, and a 4-element micro-array in Kislovodsk. The identification accuracy of the neural networks is >90% for most of the stations that we tested. The neural network module was integrated into the Intelligent Monitoring System (*IMS*), and it was applied to the 3-component IRIS/IDA data under simulated operational conditions. The result was a reduction in the number false-alarms produced by the automated processing and interpretation system by about 60%.

Table of Contents

I. Introduction

1.1 Background

The Intelligent Monitoring System (*IMS*) is one of several related systems that SAIC and it subcontractors (primarily Inference Corporation) have developed for automated and interactive analysis of data from a network of seismic stations to detect and locate seismic events. It has been operating nearly continuously since 1989 while evolving through several increasingly capable versions. The first version was used to detect and locate regional events recorded by two 25-element arrays in Norway; NORESS and ARCESS [*Bache et al.*, 1990*a*,*b*]. The second version was extended to detect and locate all seismic events recorded by an arbitrary network [*Bache et al.*, 1991]. This version was installed for operation in November, 1990, with data from NORESS, ARCESS, and a 16-element array in Finland (FINESA). In March 1991, a 25-element array in Germany (GERESS) was added to "*IMS* Version 2."

Data from two 3-component stations in Poland (KSP and SFP) were added to *IMS* Version 2 between June and December, 1991, and this motivated the development of our neural network for identifying initial wave type [*Patnaik and Sereno*, 1991]. *IMS* Version 2 used simple rules that were based on a few polarization attributes to identify *P* and *S* waves recorded by 3-component stations. *Suteau-Henson* [1991] used a multivariate discriminant analysis to show that the identification accuracy could be increased by including several other polarization attributes, and that the optimal discriminants are station-dependent. *Patnaik and Sereno* [1991] extended this by adding other polarization and contextual attributes, and by replacing the linear multivariate method with a non-linear neural network. They concluded that the major advantages of this approach are: (1) it is easier to develop neural networks than it is to formulate rules for high-dimensional input, (2) station-specific characteristics are easy to represent, (3) neural networks are 3-7% more accurate than the linear multivariate method, and (4) neural networks are easily adapted to data from new stations. This neural network was integrated into the *Expert System for Association and Location* (**ESAL**) for operational test and evaluation in August, 1991. **ESAL** is the major knowledge-based component of *IMS* Version 2 [*Bratt et al.*, 1991].

The neural networks were tested under simulated operational conditions using data recorded by six 3-component IRIS/IDA stations in the CIS [*Patnaik et al.*, 1992]. Neural networks were trained for two of these stations, but the others had too few analyst-reviewed data for reliable training and testing. For these, average neural network weights were developed by training with data from all six stations. **ESAL** was applied twice to a 6-week IRIS/IDA data set; once with the rule-based system for initial wave-type identification, and once with the neural network. The identification accuracy was about 95% for each of the stations with individually-trained neural networks, which was 3-6% higher than it was for the rule-based system. The final result was a more accurate automated event bulletin when the neural network was used. However, even with this improvement most of the events formed by **ESAL** were either rejected or ignored by the analyst (i.e., false-alarms).

Most of the false-alarms were formed from noise detections at 3-component stations that were not recognized as such by **ESAL**. Instead, **ESAL** identified all detections at 3-component stations as either *P* or *S* phases. This problem was not encountered with data from the high-frequency

regional arrays because the phase velocity determined from *f-k* analysis could be used to reliably discriminate between signal and noise. Under our current ARPA contract, two major enhancements were made to the neural network to address the false-alarm problem. First, we extended the neural network output from two classes (*P* or *S*) to four classes (Teleseism, Regional *P*, Regional *S*, or Noise). Second, we added a new spectral representation of the horizontal-to-vertical power ratio of each arrival to the input attributes. As described in this report, the result of these enhancements was a reduction in the number of false-alarms by about 60%. We also generalized the neural network to include slowness as an input attribute for array stations. This was motivated by the addition of data from mini- and micro-arrays whose *f-k* resolution is not adequate for reliable noise screening. These include 9-element arrays in Apatity and Spitsbergen [*Mykkeltveit et al.*, 1992], and a 4-element array in Kislovodsk [*Berger et al.*, 1992].

1.2 Overview

The objective of this report is to describe *IMS* Version 3 which we define as an extension of *IMS* Version 2 to include the new 4-class neural network for initial wave-type identification. Figure 1 shows the **ESAL** implementation of the neural network software module. The signal processing component of *IMS* performs detection and feature extraction (arrival time, frequency, polarization attributes, etc.). These features are interpreted by **ESAL** in two major stages (Station and Network Processing) that are essentially independent [*Bratt et al.*, 1991]. Station Processing identifies initial wave type, forms groups of phases that appear to be from the same event, identifies as many of these phases as possible, and computes single-station event locations when there are adequate data. Network Processing uses results from Station Processing to associate as many phases as possible and to form all plausible event solutions. Detailed descriptions of **ESAL**'s Network Processing are given by *Bratt et al.* [1991] and *Bache et al.* [1993].

The neural network software module was implemented as one of two options for initial wave-type identification in **ESAL** Station Processing. The other option is the current rule-based system. The neural network option will default to the rule-based system if a neural network was not trained for a particular station, or if any of the input attributes are not available. The current implementation includes dominant period, seven polarization attributes, two contextual attributes (e.g., measurements related to a group of arrivals), horizontal-to-vertical power ratios at five different frequencies, and *f-k* slowness for array stations. The output of the neural network is the initial wave type (Teleseism, Regional *P*, Regional *S*, or Noise) and a measure of confidence. The initial wave type is used in *Phase Grouping*, but the confidence measure is currently not used.

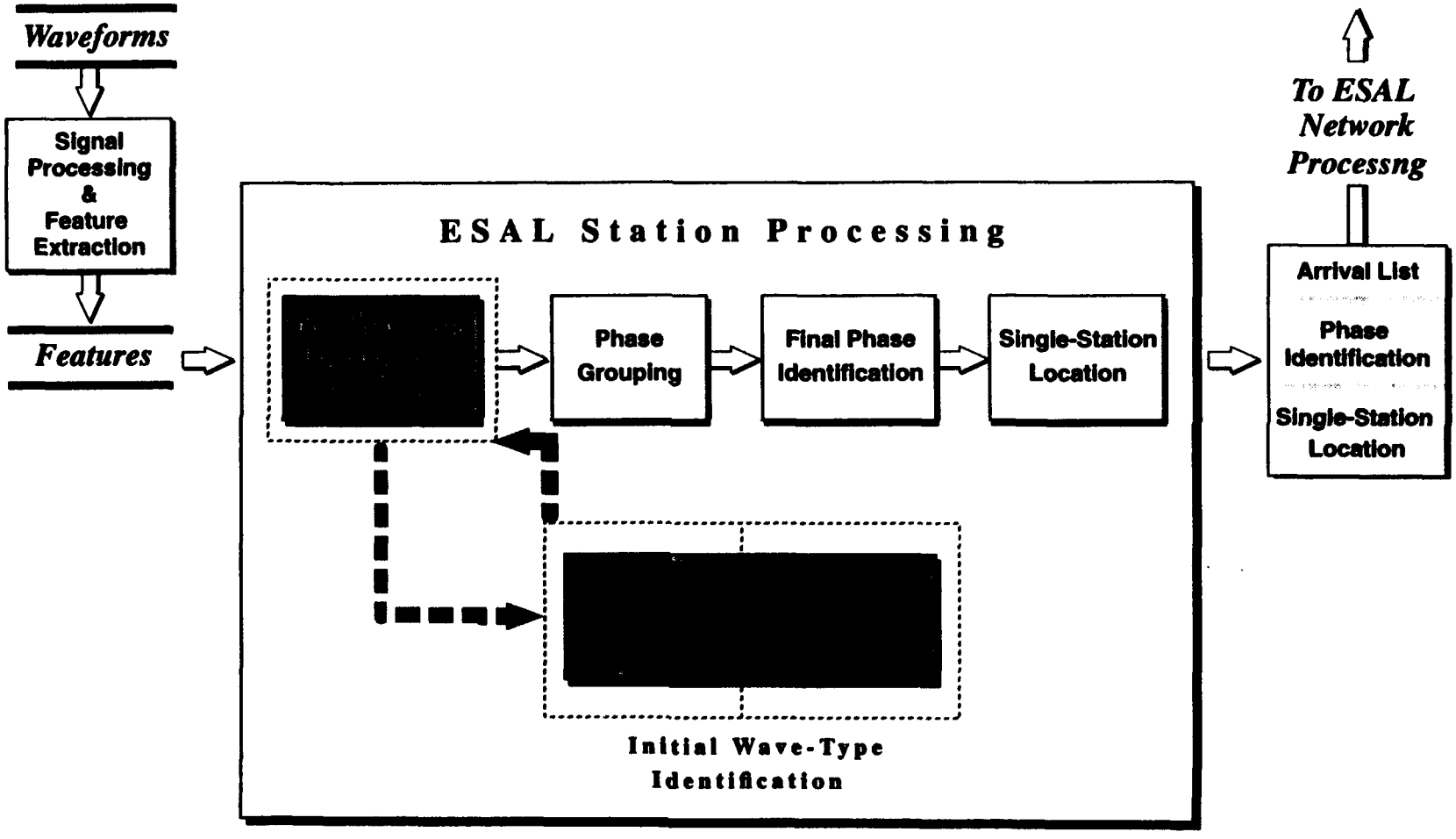


Figure 1. This schematic illustration shows the **ESAL** implementation of the 4-class neural network software module for initial wave-type identification.

II. Neural Network for Initial Wave-Type Identification

This section describes the input data, general architecture, and training procedures for the neural networks for initial wave-type identification.

2.1 Input Data

The input data include dominant period, polarization attributes, contextual attributes, a spectral representation of the horizontal-to-vertical power ratio, and the slowness determined by *f-k* analysis for array stations. The method used for polarization analysis was developed by *Jurkevics* [1988], and its *IMS* implementation is described by *Bache et al.* [1990*b*]. The polarization ellipsoid is computed within overlapping time windows by solving the eigenvalue problem for the covariance matrix. The covariance matrices are computed in the time domain for several frequency bands, and then normalized and averaged to obtain a wide-band estimate for each of the overlapping windows. *P*-type attributes are calculated from the window with the maximum rectilinearity, and *S*-type attributes are calculated from the window with the maximum 3-component amplitude.

The seven polarization attributes used by the neural network are described in detail by *Patnaik and Sereno* [1991] and by *Swanger et al.* [1993]. These are rectilinearity (*rect*), planarity (*plans*), horizontal-to-vertical power ratio measured at the time of maximum rectilinearity (*hvratp*) and at the time of the maximum 3-component amplitude (*hvrat*), maximum-to-minimum horizontal amplitude ratio (*hmxmn*), the short-axis incidence angle (*inang3*), and the long-axis incidence angle (*inang1*). The neural network inputs are scaled to a small numerical range near ±1 by dividing the incidence angles by 90 degrees, and taking the common logarithm of the amplitude and power ratios. Pre-processing is not required for the other polarization attributes, as their numerical values are already limited to acceptable ranges. Examples of dominant period and the seven polarization attributes are given in Figures 2-9 for each of the 3-component IRIS stations in the CIS for a one-week data set recorded in July, 1991.

Two attributes are used to parameterize contextual information about a group of arrivals. One of these is the difference between the number of arrivals before and after the arrival in question within a fixed time window. For example, regional *P* phases are more likely to have arrivals after them than before but the opposite is more likely for regional *S* phases. The other contextual attribute is the mean time difference between the arrival in question and arrivals before and after it within the same fixed time window [*Patnaik and Sereno*, 1991]. Examples of these attributes are given in Figures 10-11 for the six IRIS stations. The time windows can be determined empirically for each station, but we used 60 s for each of the IRIS stations because of the limited amount of analyst-reviewed training data. The first contextual attribute is scaled to a small range near ±1 by dividing by 10, and the second contextual attribute is divided by 100 s.

The broadband horizontal-to-vertical power ratio is one of the most useful polarization attributes for identifying initial wave type, so we extended it by parameterizing its frequency dependence. We calculated an average horizontal component from the N-S and E-W components using $(N\text{-}S^2 + E\text{-}W^2)^{1/2}$. The horizontal and vertical components were filtered in 5 one-octave bands centered at 0.25, 0.5, 1.0, 2.0, and 4.0 Hz. The peak horizontal and vertical amplitudes were measured

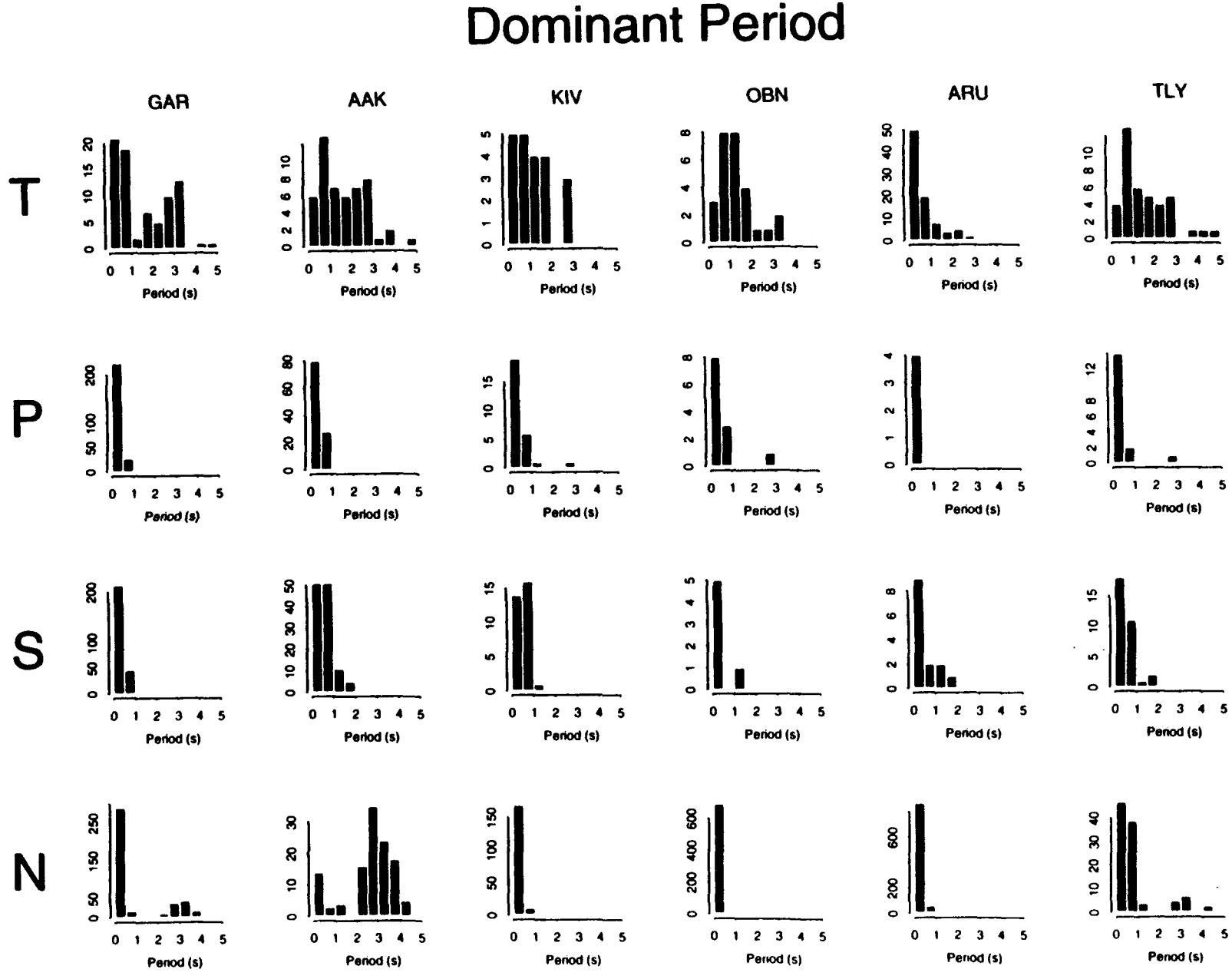


Figure 2. Histograms of dominant period are plotted for teleseisms, regional *P*, regional *S* and noise detections for six IRIS/IDA stations in the CIS.

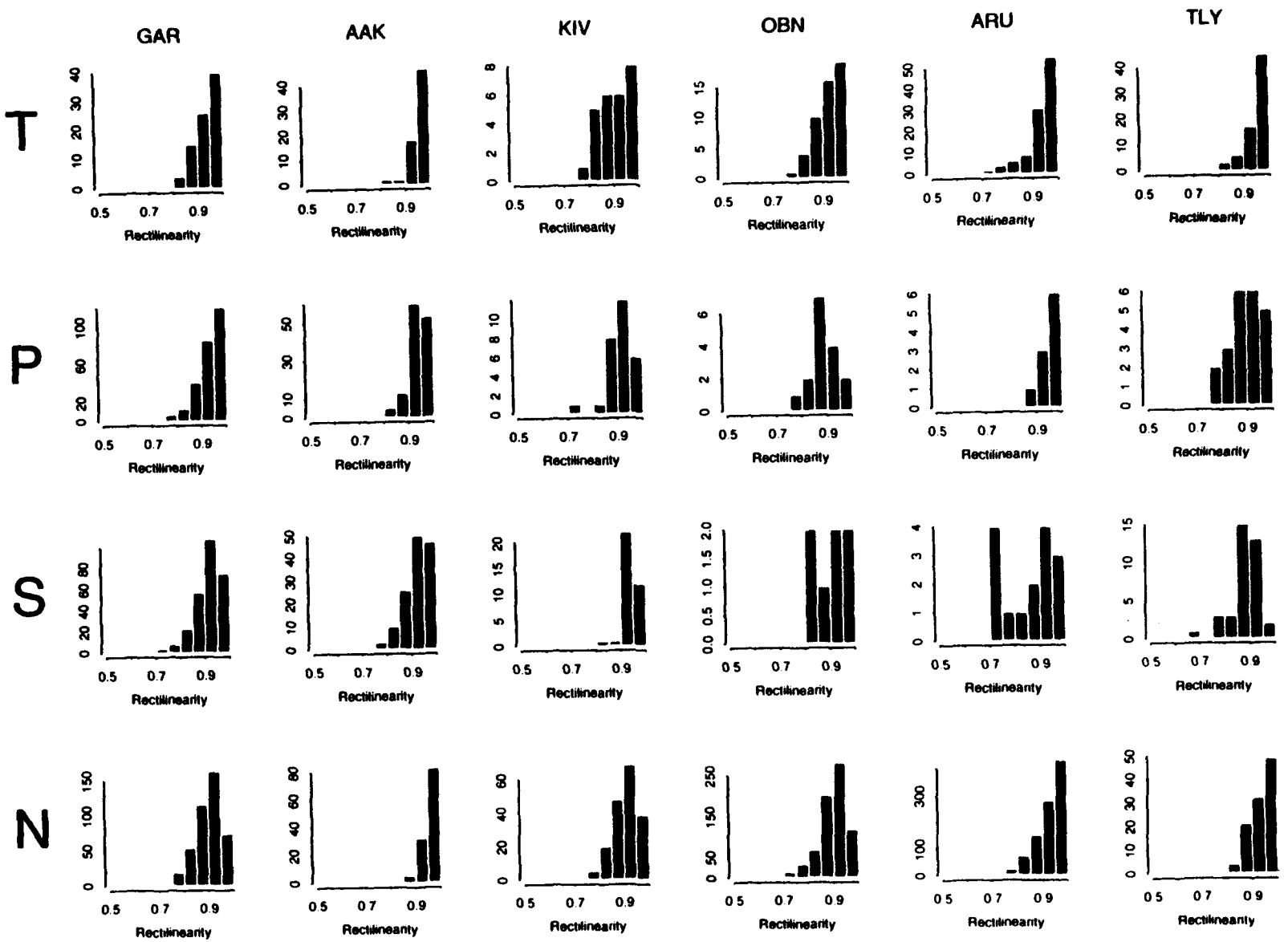


Figure 3. Histograms of rectilinearity are plotted for teleseisms, regional *P*, regional *S* and noise detections for six IRIS/IDA stations in the CIS.

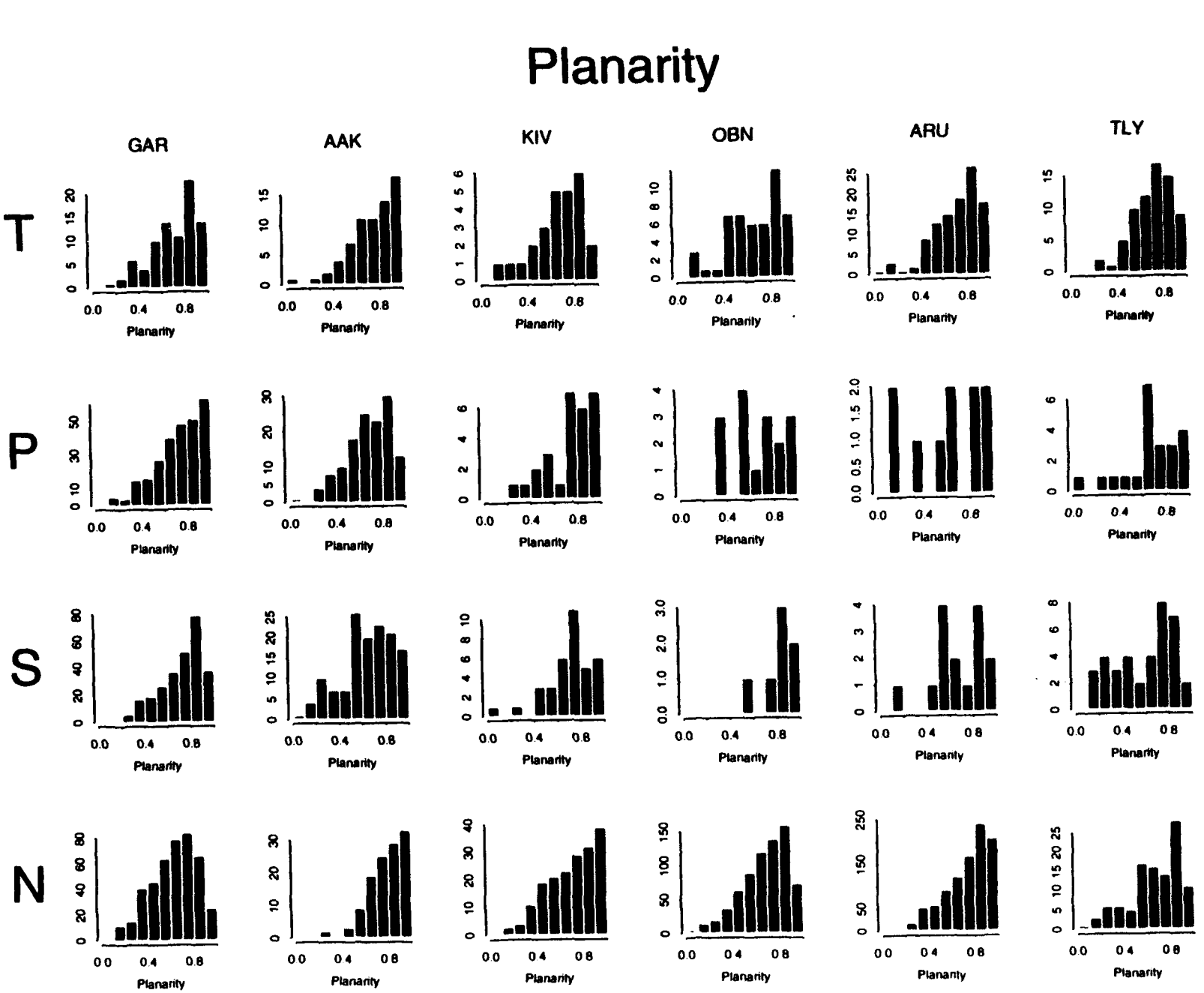


Figure 4. Histograms of planarity are plotted for teleseisms, regional *P*, regional *S* and noise detections for six IRIS/IDA stations in the CIS.

Horizontal-to-Vertical Power Ratio (S-Type)

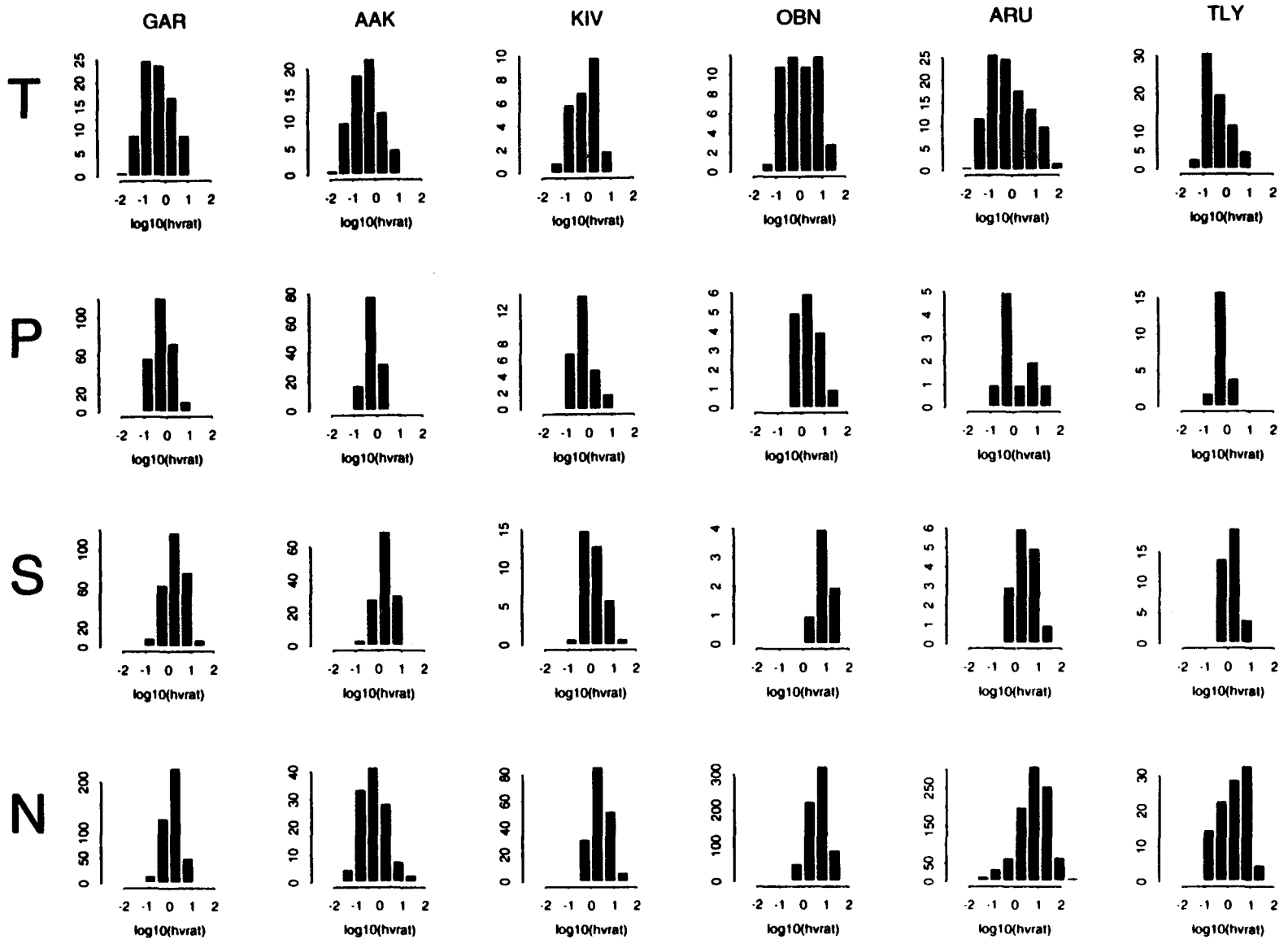


Figure 5. Histograms of horizontal-to-vertical power ratio at the time of maximum 3-component amplitude are plotted for teleseisms, regional *P*, regional *S* and noise detections for six IRIS/IDA stations in the CIS.

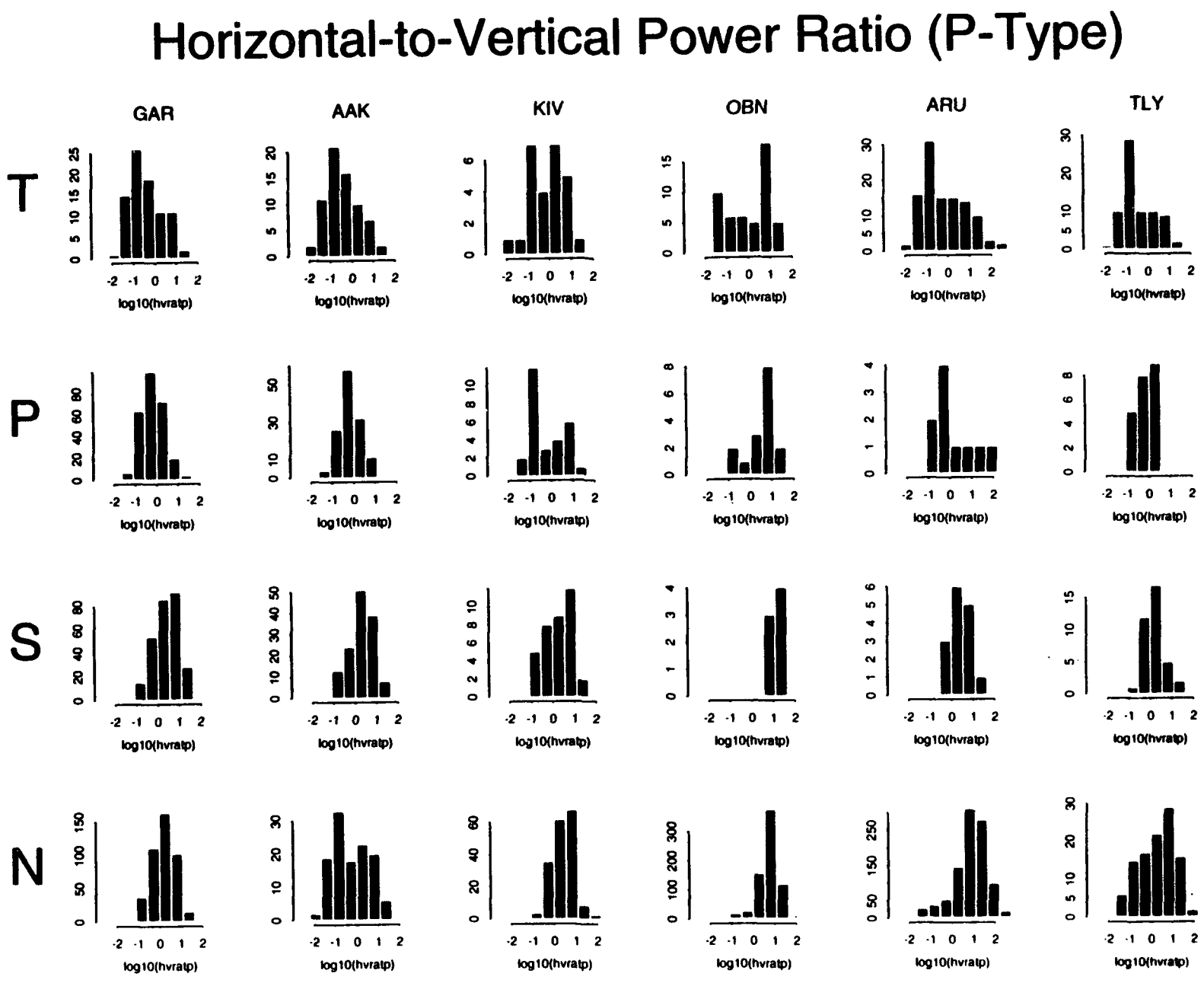


Figure 6. Histograms of horizontal-to-vertical power ratio at the time of maximum rectilinearity are plotted for teleseisms, regional *P*, regional *S* and noise detections for six IRIS/IDA stations in the CIS.

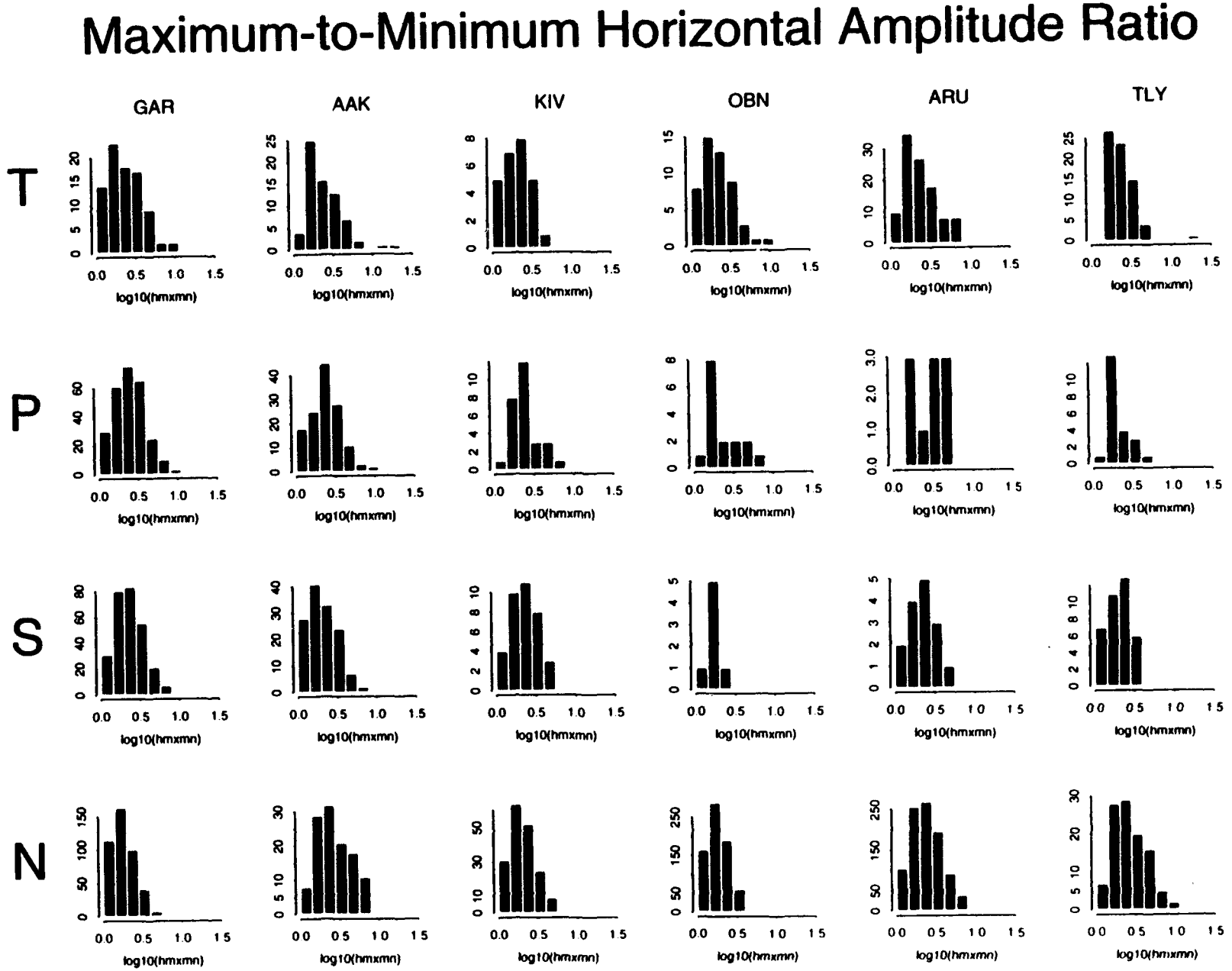


Figure 7. Histograms of maximum-to-minimum horizontal amplitude ratio are plotted for teleseisms, regional *P*, regional *S* and noise detections for six IRIS/IDA stations in the CIS.

Short-Axis Incidence Angle

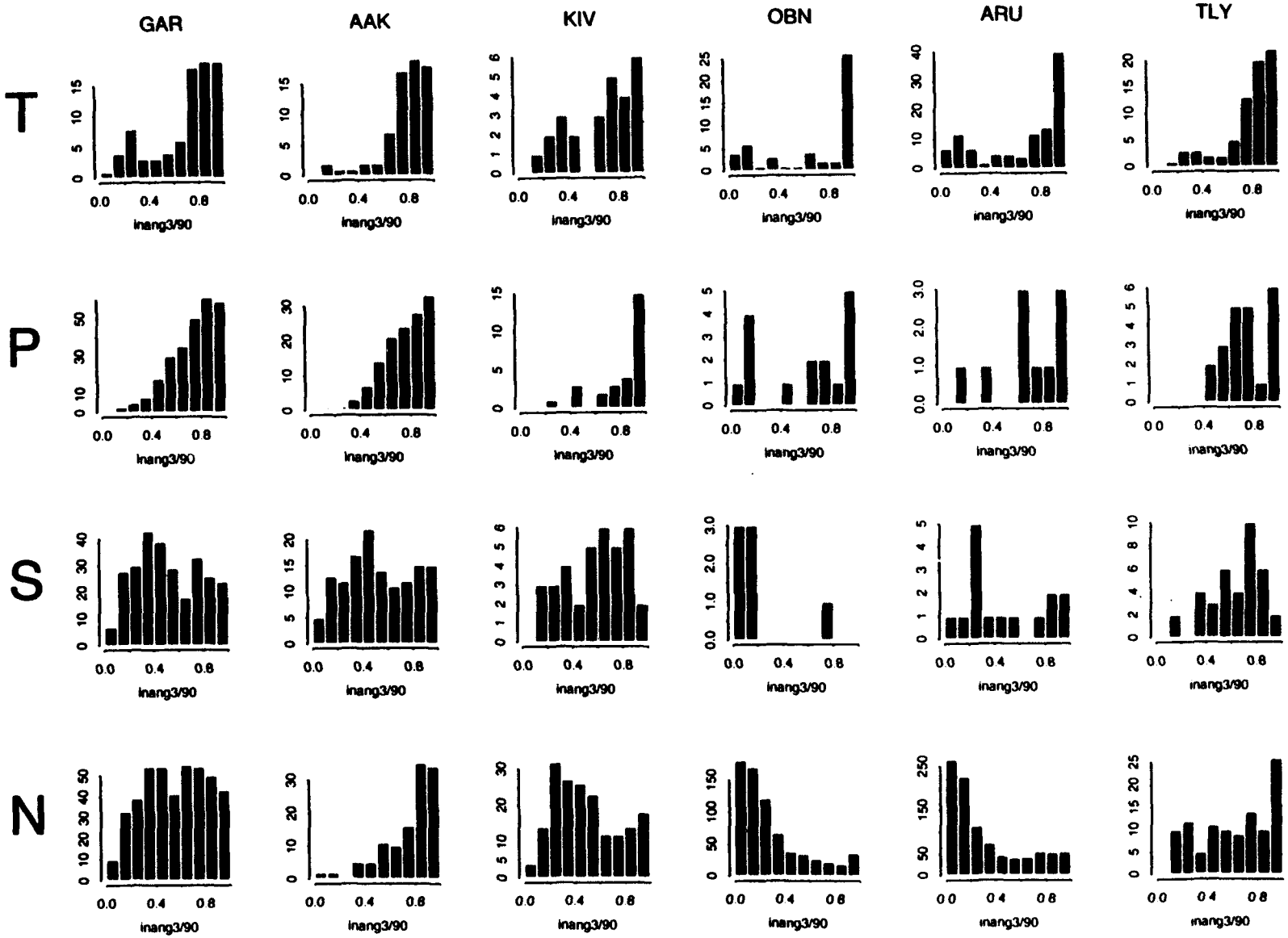


Figure 8. Histograms of short-axis incidence angle are plotted for teleseisms, regional *P*, regional *S* and noise detections for six IRIS/IDA stations in the CIS.

Long-Axis Incidence Angle

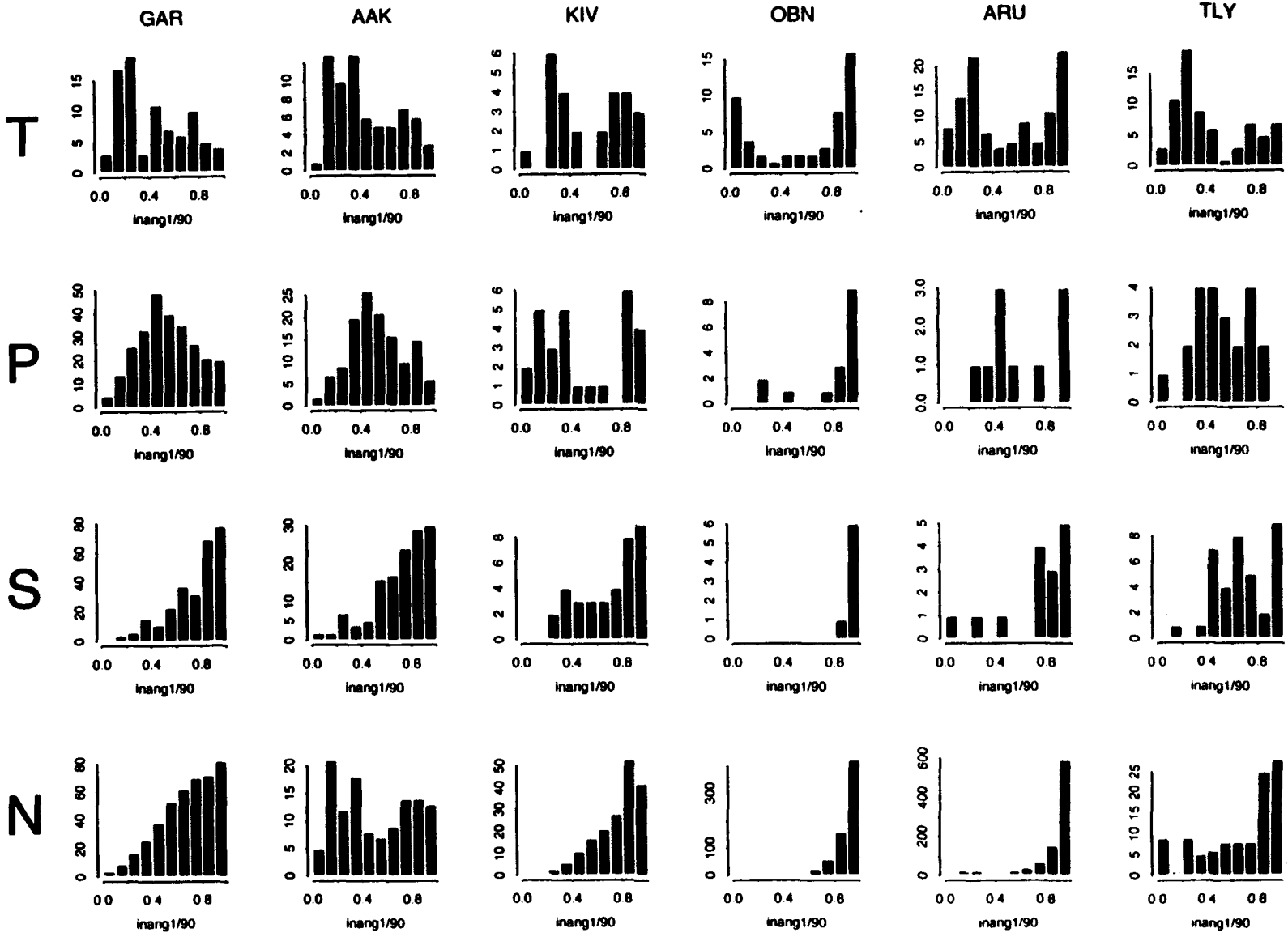


Figure 9. Histograms of long-axis incidence angle are plotted for teleseisms, regional *P*, regional *S* and noise detections for six IRIS/IDA stations in the CIS.

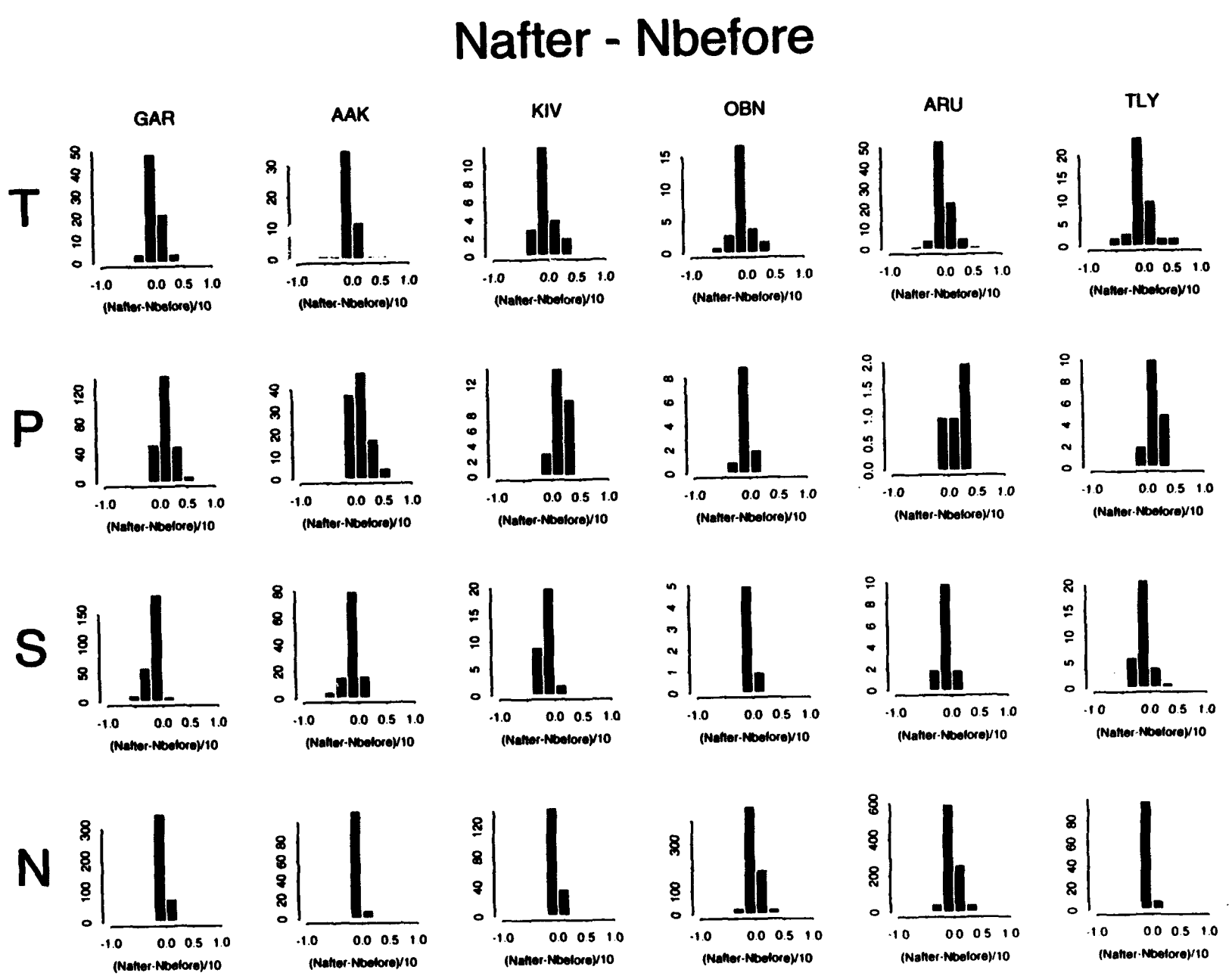


Figure 10. Histograms of the number of detections that arrive after and before the detection in question within a fixed time window of 60-s are plotted for teleseisms, regional *P*, regional *S* and noise detections for six IRIS/IDA stations in the CIS.

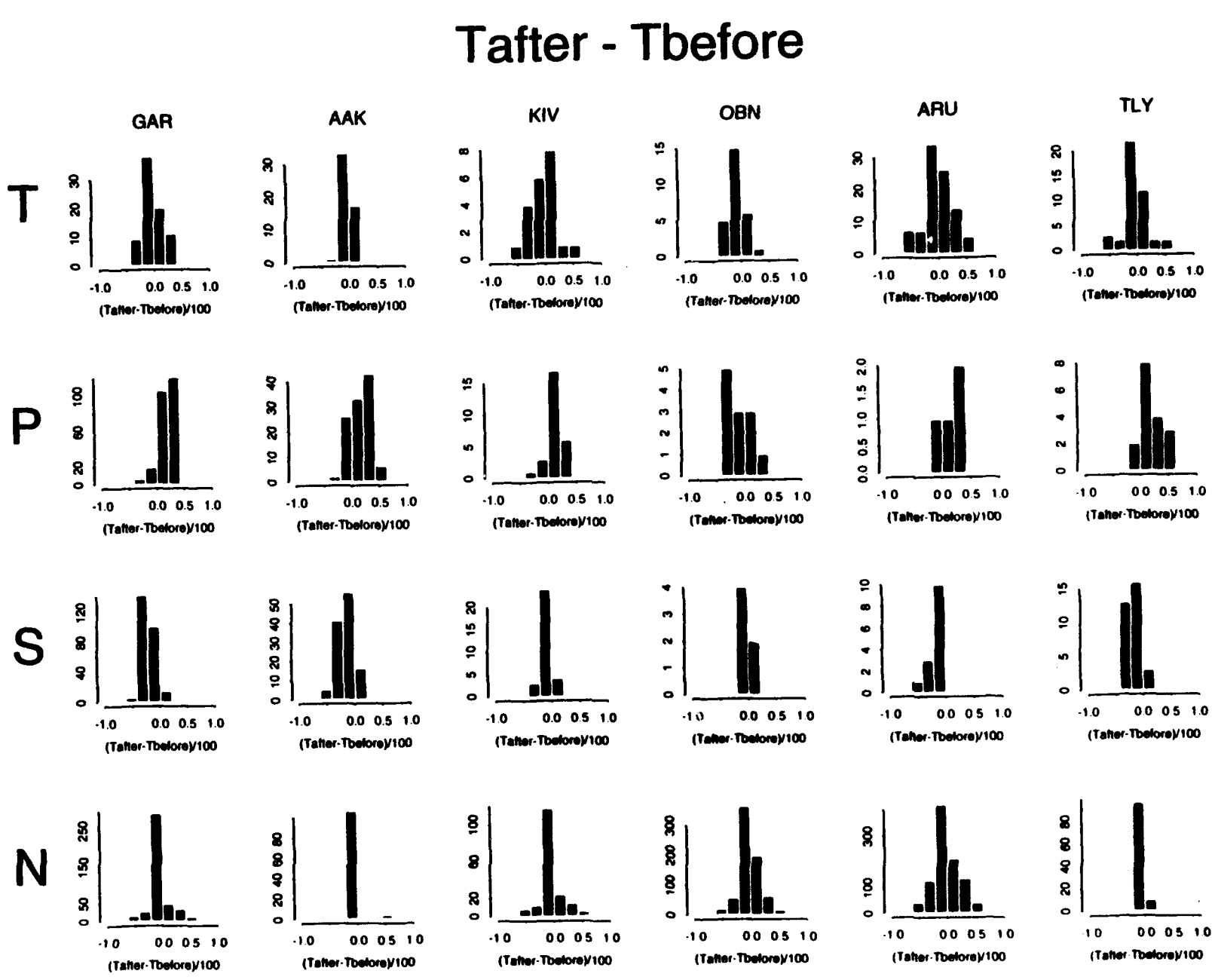


Figure 11. Histograms of the difference between the average arrival times after and before the detection in question within a fixed time window of 60-s are plotted for teleseisms, regional *P*, regional *S* and noise detections for six IRIS/IDA stations in the CIS.

using a time window that started 4 s before the detection time with a duration of 10-s for the three lowest-frequency bands, and 8 s for the two highest-frequency bands. The horizontal-to-vertical power ratio is defined as $H^2/2V^2$ where H is the amplitude on the horizontal component and V is the amplitude on the vertical component. Figures 12-16 show the frequency-dependent horizontal-to-vertical power ratios for each of the six IRIS stations in the CIS. The logarithms of these ratios are used as input to the neural networks.

Slowness can be estimated accurately enough from *f-k* analysis to enable nearly perfect initial wave-type identification from data recorded by NORESS-type arrays. For example, the left side of Figure 17 shows *f-k* slowness as a function of initial wave type for data recorded at ARCESS between July and September, 1991. More recently, data have been added to *IMS* from mini-arrays in Apatity and Spitsbergen and a 4-element micro-array in Kislovodsk. The *f-k* resolution is significantly lower for these arrays (Figure 18). The last three columns in Figure 17 show *f-k* slowness as a function of wave type for these smaller arrays. The data from Apatity were recorded between January and March, 1993, the data from Spitsbergen were recorded between April and June, 1993, and the data from Kislovodsk were recorded during a two-week period in October, 1992. Clearly, the *f-k* slowness alone is not adequate for reliable identification of initial wave type for these smaller arrays. Therefore, we generalized the neural network for small-aperture arrays by including *f-k* slowness as an input attribute. Slowness in s/km is confined to a small numerical range near ± 1, so this attribute does not require pre-processing.

2.2 Neural Network Architecture

The four-class initial wave-type identification problem is solved in three stages, and each stage consists of a neural network that solves a two-class problem. This is illustrated in Figure 19 which is a schematic representation of our approach to the four-class problem. The dark ovals are the final four output classes: Teleseisms, Regional *P*, Regional *S*, and Noise. The Stage 1 neural network separates signal and noise. The Stage 2 neural network is applied to all signal detections from Stage 1, and it separates Regional *S* phases from Regional *P* phases and Teleseisms. The Stage 3 neural network is applied to all Teleseisms and Regional *P* phases from Stage 2, and it separates these into two distinct classes. The input at each stage consists of the signal features described in the previous section.

The general architecture of each of the two-class neural networks is described by *Patnaik and Sereno* [1991], and it is illustrated schematically in Figure 20. They consist of three-layers: an input layer, a middle (or hidden) layer, and an output layer. The input layer consists of 15 or 16 nodes (15 for 3-component stations and 16 for array stations) corresponding to the attributes described in the previous section. The number of nodes in the middle layer is determined empirically for each station as described by *Patnaik and Sereno* [1991]. Although this can be station-dependent, we found that 6 nodes in the middle layer generally provided satisfactory results for all stations that we tested. There are two nodes in the output layer corresponding to the classes for each of the three neural networks in Figure 19.

Networks of the type shown in Figure 20 are called *layered feedforward* neural networks by *Rumelhart et al.* [1986]. In these networks, the input nodes are the bottom layer and the output nodes are the top layer. There can be any number of hidden layers in between (in our case there is

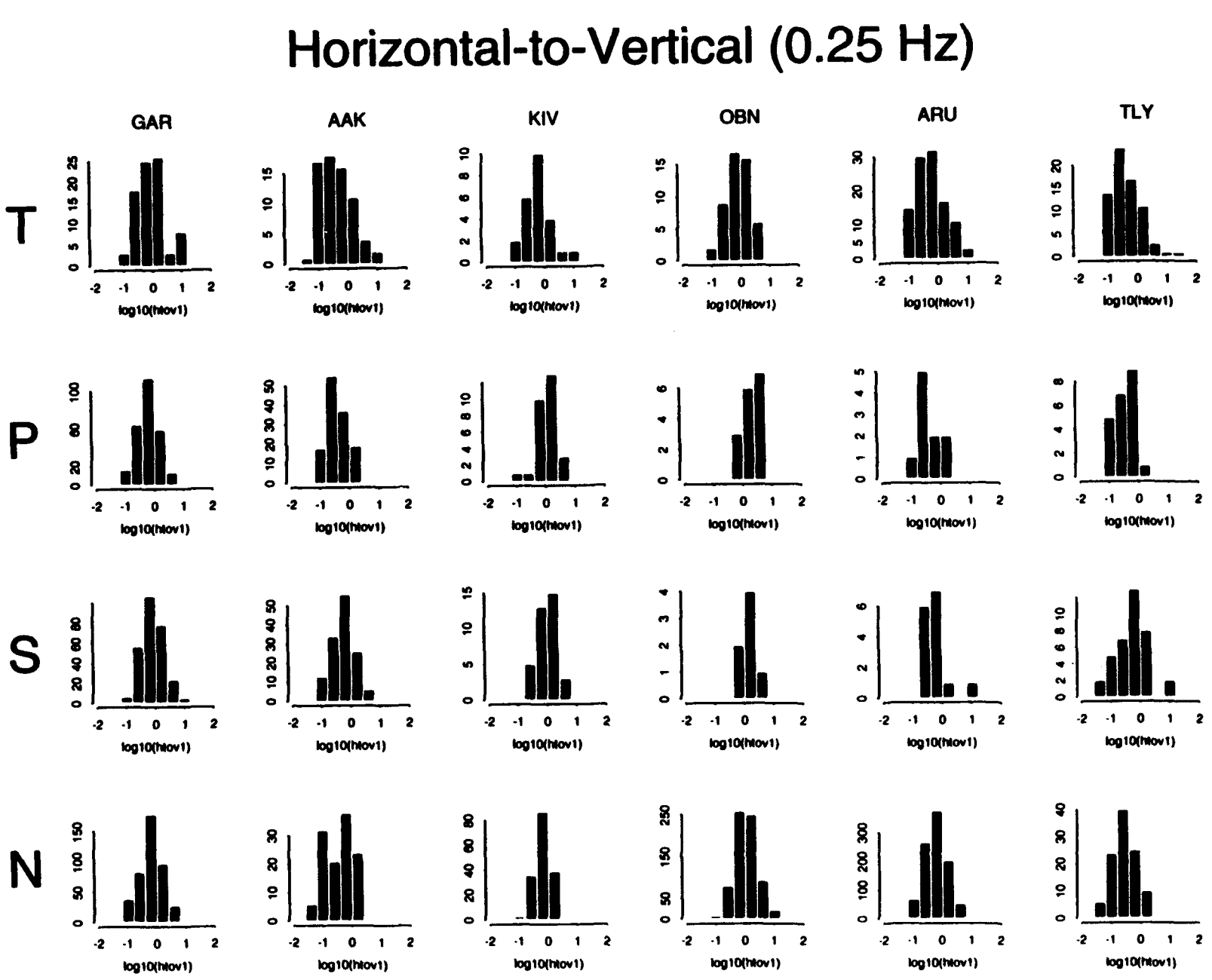


Figure 12. Histograms of the horizontal-to-vertical power ratio at a center frequency of 0.25 Hz are plotted for teleseisms, regional *P*, regional *S* and noise detections for six IRIS/IDA stations in the CIS.

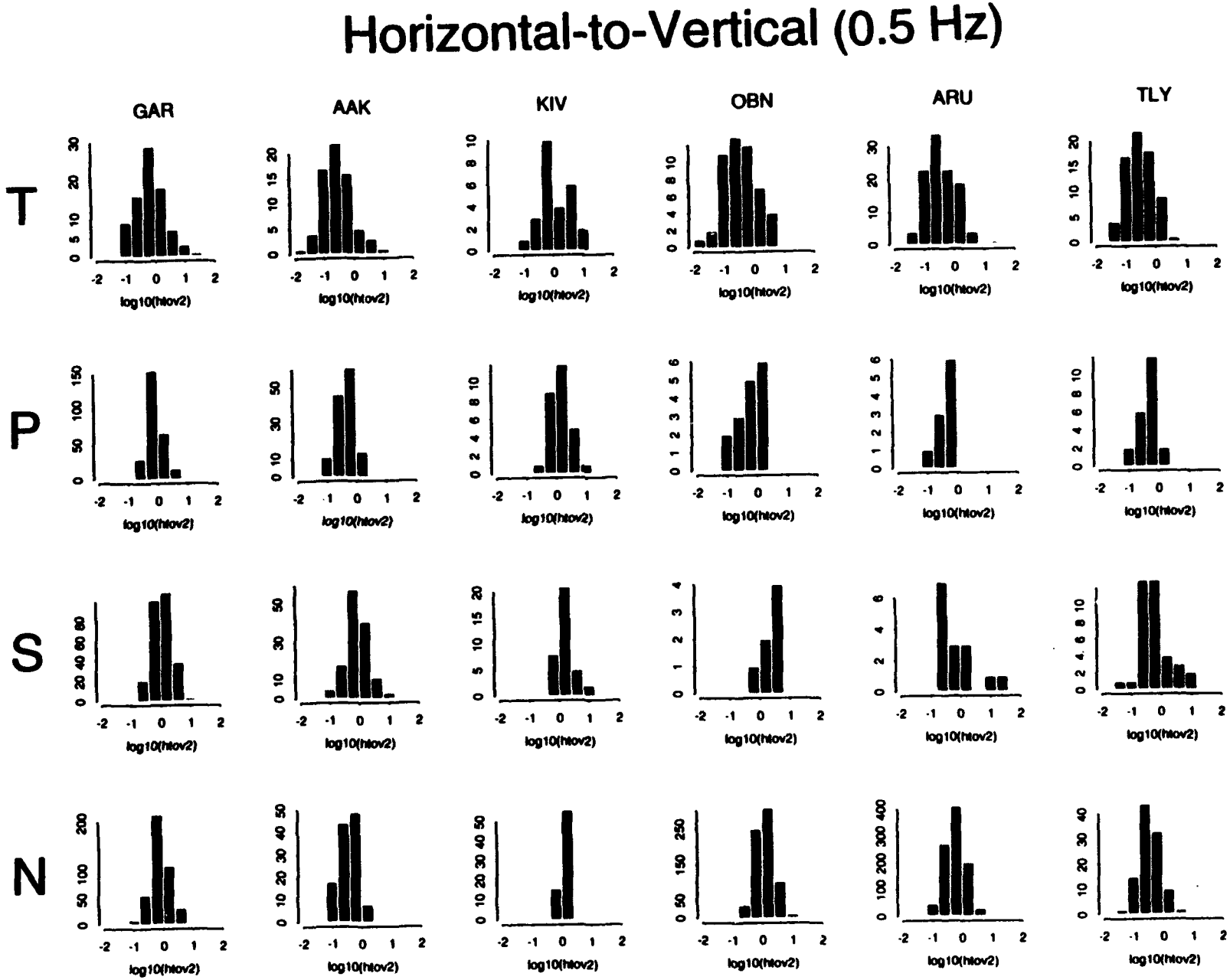


Figure 13. Histograms of the horizontal-to-vertical power ratio at a center frequency of 0.5 Hz are plotted for teleseisms, regional *P*, regional *S* and noise detections for six IRIS/IDA stations in the CIS.

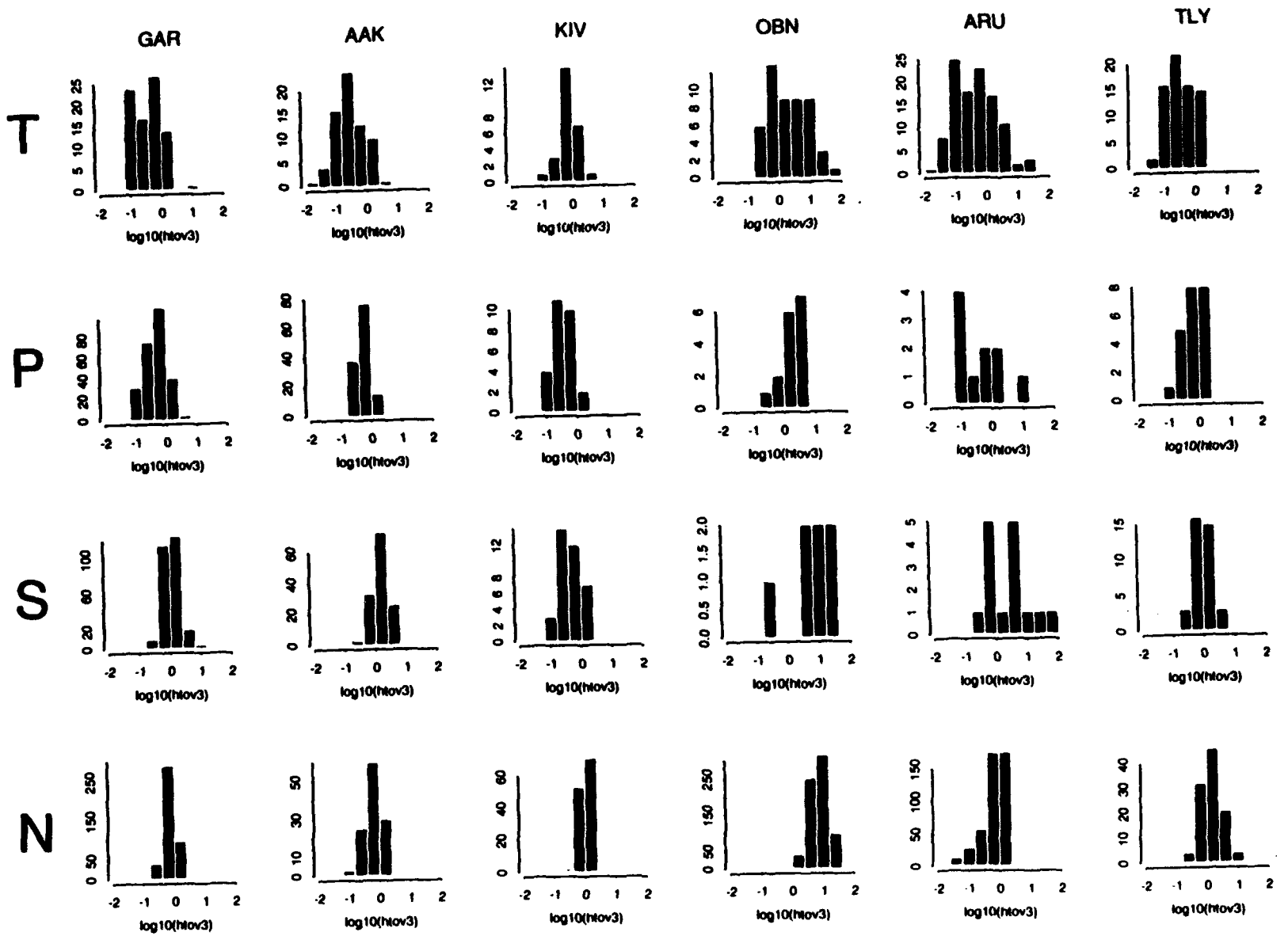


Figure 14. Histograms of the horizontal-to-vertical power ratio at a center frequency of1.0 Hz are plotted for teleseisms, regional *P*, regional *S* and noise detections for six IRIS/IDA stations in the CIS.

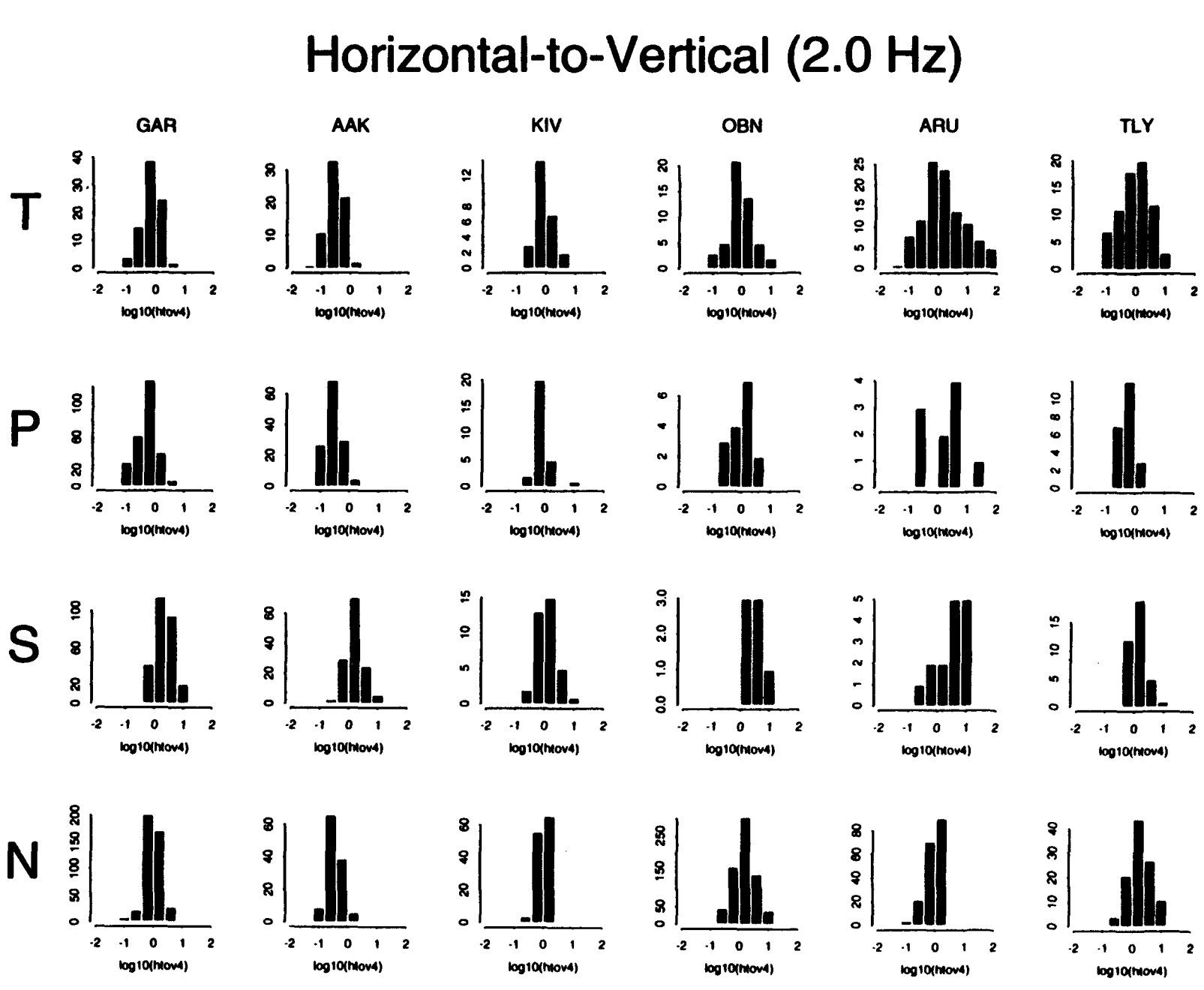


Figure 15. Histograms of the horizontal-to-vertical power ratio at a center frequency of 2.0 Hz are plotted for teleseisms, regional *P*, regional *S* and noise detections for six IRIS/IDA stations in the CIS.

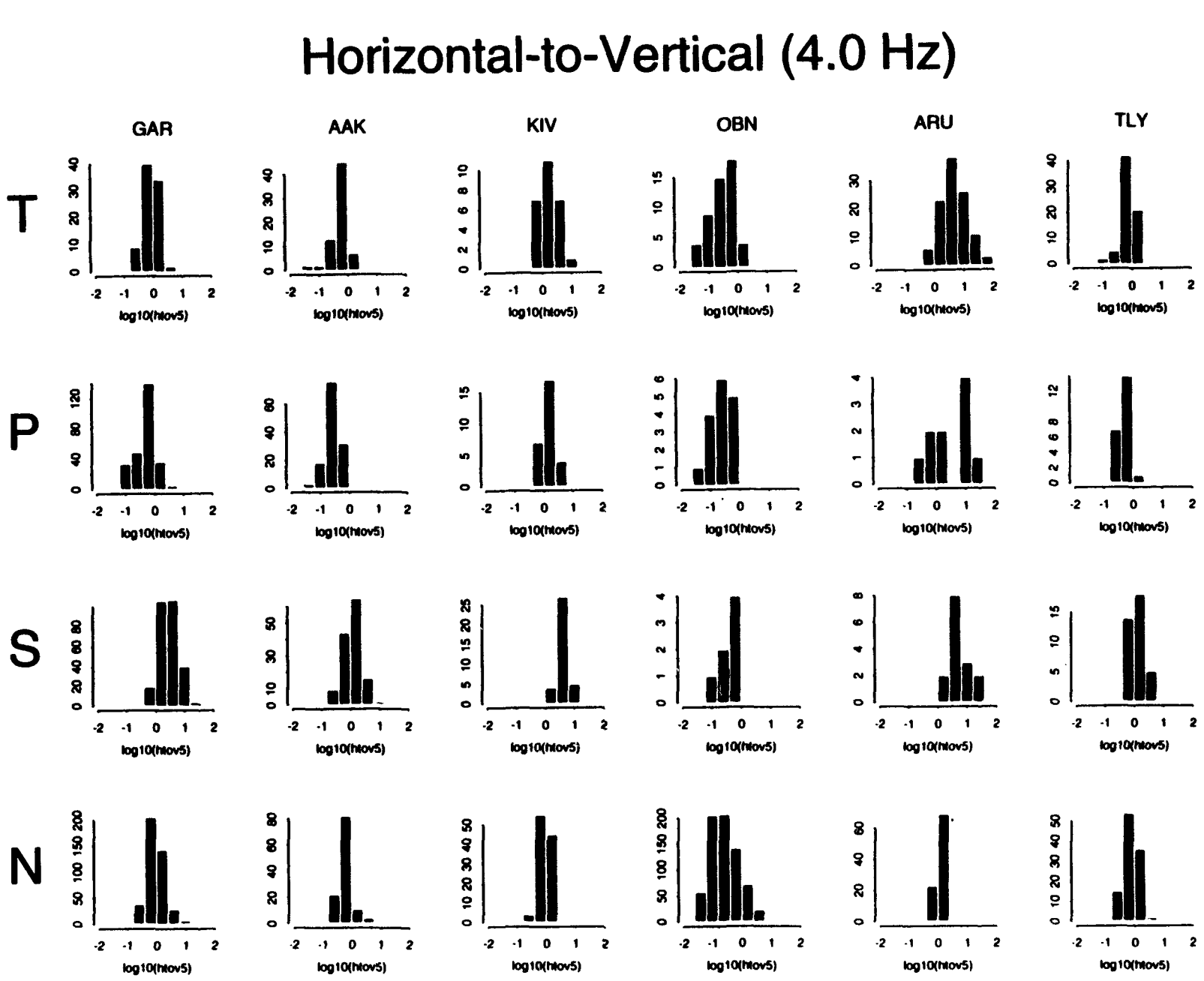


Figure 16. Histograms of the horizontal-to-vertical power ratio at a center frequency of 4.0 Hz are plotted for teleseisms, regional *P*, regional *S* and noise detections for six IRIS/IDA stations in the CIS.

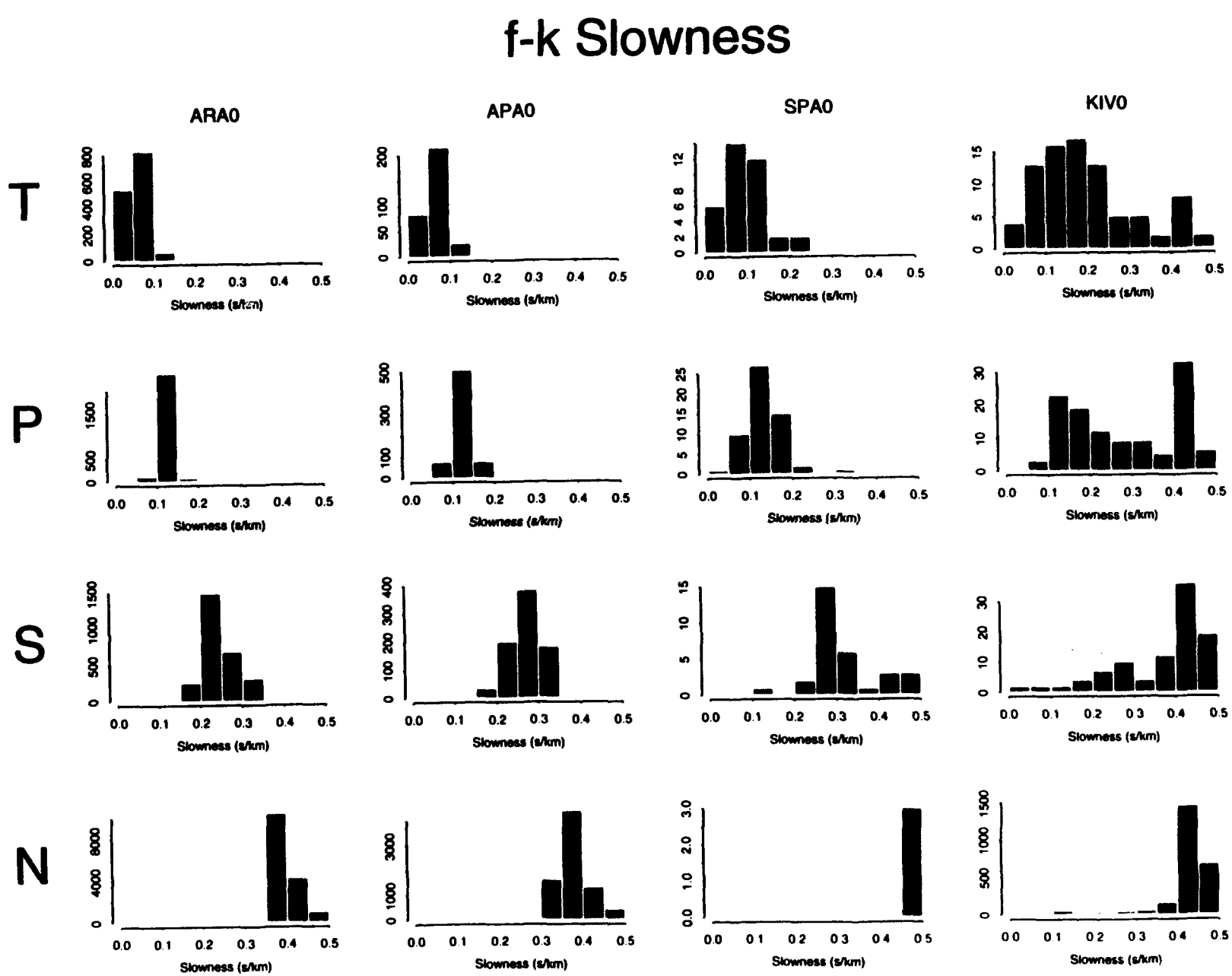


Figure 17. Histograms of *f-k* slowness are plotted for teleseisms, regional *P*, regional *S* and noise detections at ARCESS (ARA0), Apatity (APA0), Spitsbergen (SPA0), and Kislovodsk (KIV0).

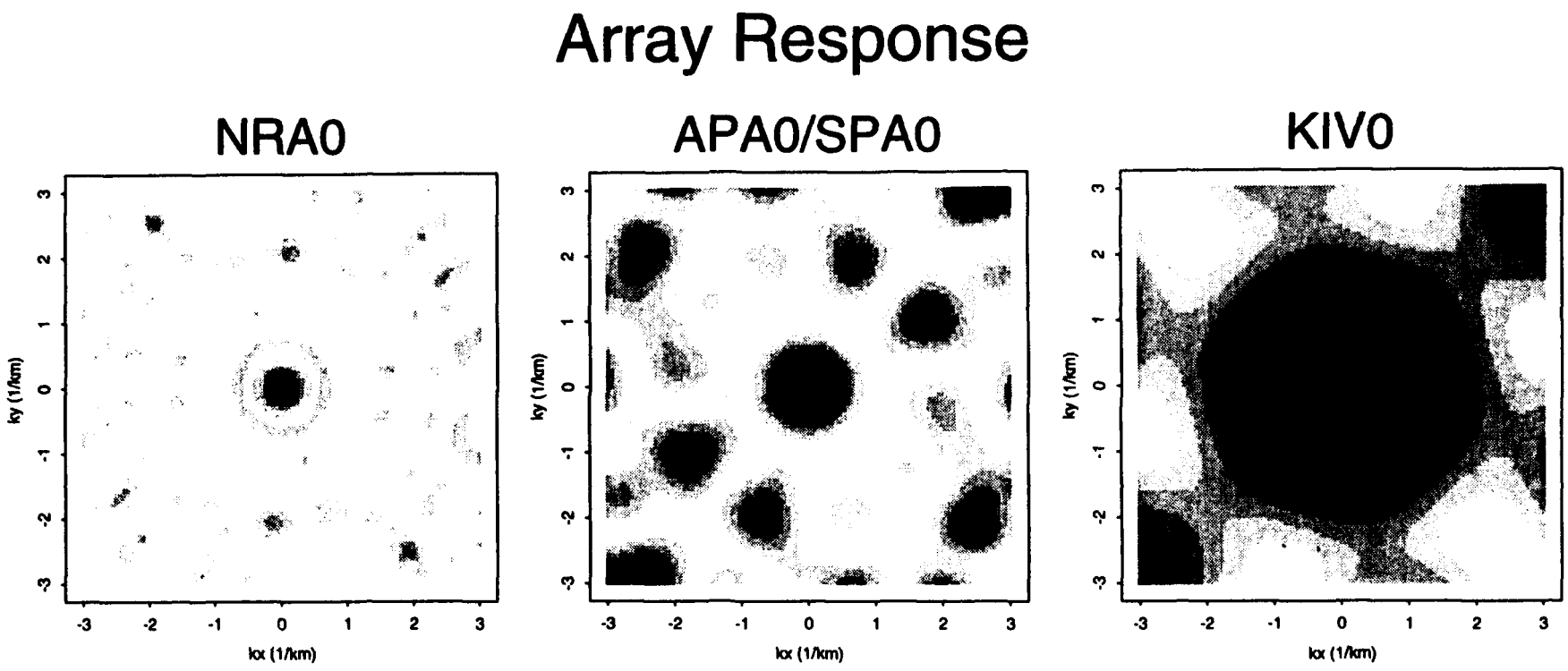


Figure 18. The array response to an infinite-velocity plane wave is plotted for the 25-element NORESS array, the 9-element mini-arrays in Spitsbergen and Apatity, and the 4-element micro-array in Kislovodsk.

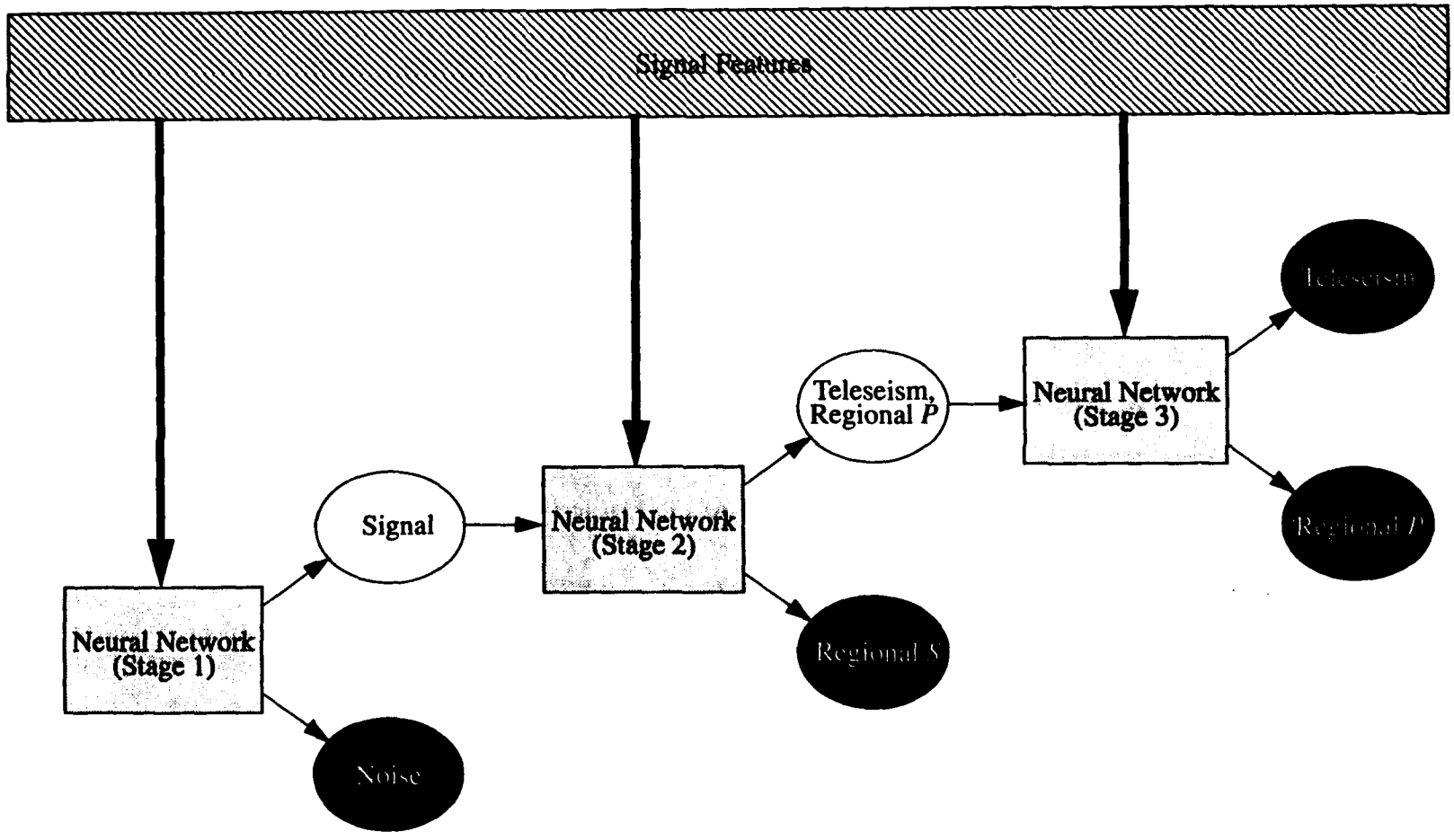


Figure 19. A schematic representation of our neural network approach to solving the 4-class initial wave-type identification problem is shown.

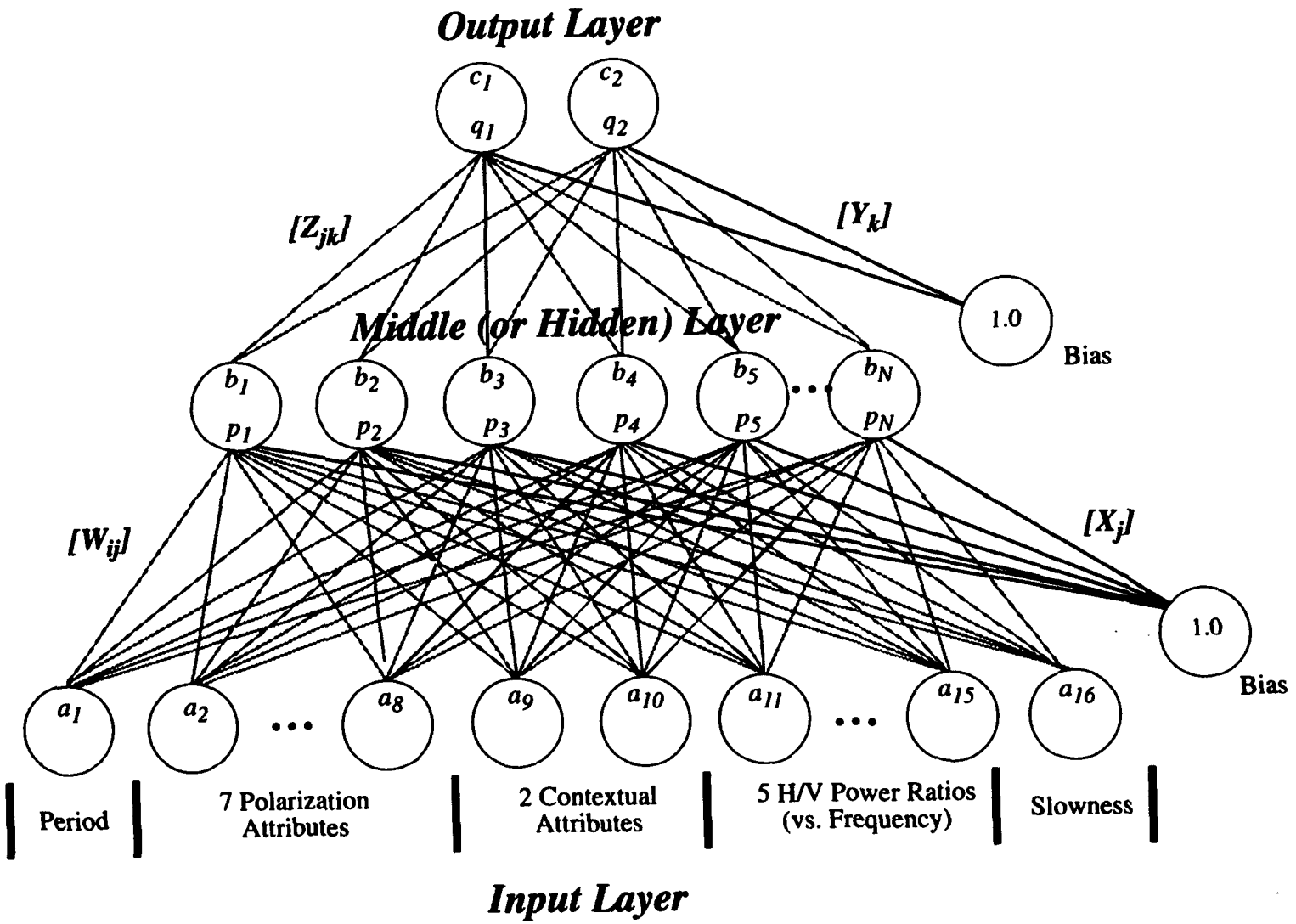


Figure 20. A three-layer, feedforward neural network is illustrated.

only one), but any node in a hidden layer must send its output to higher layers and receive its input from lower layers. The input to each node in the middle layer, p_j, is a linear combination of the input attributes, a_i, and a bias term to add translation to the functions modeled by the network. That is, we define the input to each node in the middle layer as:

$$p_j = \sum_{i=1}^{M} a_i W_{ij} + X_j \tag{1}$$

where W_{ij} and X_j are station-specific weights to be determined during network training (the X_j are associated with the bias term shown in Figure 20), and M is the number of input nodes (i.e., either 15 or 16). The output of each node in the middle layer, b_j, is calculated by applying a nonlinear activation function to its input. We use:

$$b_j = f(p_j) = \frac{1}{(1 + e^{-p_j})} \tag{2}$$

which is called *semilinear* by *Rumelhart, et al.* [1986]. They define a semilinear activation function as one in which the output is a non-decreasing and differentiable function of the net total input, p_j. This function limits the range of the output from zero to one. The input to each node in the output layer, q_k, is a linear combination of the outputs from the nodes in the middle layer:

$$q_k = \sum_{j=1}^{N} b_j Z_{jk} + Y_k \tag{3}$$

where Z_{jk} and Y_k are station-specific weights to be determined during network training, and N is the number of nodes in the middle layer. Finally, the output node activations, c_k, are calculated by applying the nonlinear activation function to q_k:

$$c_k = f(q_k) = \frac{1}{(1 + e^{-q_k})} \tag{4}$$

The output class determined by the neural network is the one with the highest node activation, c_k. *Patnaik and Sereno* [1991] developed an empirical confidence scale based on this node activation, but it is not used in the current implementation in *IMS* Version 3.0.

2.3. Neural Network Training

A backpropagation algorithm that uses signals with known initial wave types is used for training the neural networks. We assume that the wave types assigned by an analyst are accurate (i.e., *ground-truth*). Network training consists of determining the weights: W_{ij}, X_j, Z_{jk}, and Y_k. After initializing these to random values, the training is conducted in two major steps. First, each train-

ing pattern is propagated in a forward direction through all layers using equations (1)-(4) to calculate the network's output activations. The error is calculated as the difference between these and the desired outputs. Second, these errors are propagated backward through the layers, and the weights are changed according to the *generalized delta rule* for a feedforward network of semilinear nodes. The weights are changed with each new presentation of a training pattern, and patterns are presented until a convergence criterion related to the change in the sum-squared error is satisfied.

The generalized delta rule for updating the weights is based on a gradient decent in the sum-squared error, E. This error is defined as:

$$E = \frac{1}{2}\sum_{k=1}^{2}(t_k - c_k)^2 \tag{5}$$

where t_k is the ground-truth (either 0.01 or 0.99 in this case), and c_k is the network's output node activation defined by (4). The summation is over the two nodes in the output layer. *Rumelhart et al.* [1986] derive the following equations for the weight updates:

$$\Delta Z_{jk}(n+1) = \eta\delta_k b_j + \alpha\Delta Z_{jk}(n) \tag{6}$$

$$\Delta Y_k(n+1) = \eta\delta_k + \alpha\Delta Y_k(n) \tag{7}$$

$$\Delta W_{ij}(n+1) = \eta\delta_j a_i + \alpha\Delta W_{ij}(n) \tag{8}$$

$$\Delta X_j(n+1) = \eta\delta_j + \alpha\Delta X_j(n) \tag{9}$$

where n is the presentation number, η is the *learning rate*, α is the *momentum*, and δ_k and δ_j are defined below. The first term in each of these equations is derived from the gradient descent in the sum-squared error, and the second term determines the effect of past weight changes on the current weight changes (i.e., the momentum). The equation for Y_k is the same as the one for Z_{jk} except that the input to Y_k is the bias node and it is always set to 1.0. An analogous relationship exists between W_{ij} and X_j. The two parameters, η and α, govern the rate at which the learning takes place, and these are determined empirically.

The functions δ_k and δ_j in (6)-(9) are defined as:

$$\delta_k = (t_k - c_k)f'(q_k) \tag{10}$$

$$\delta_j = f'(p_j)\sum_k \delta_k Z_{jk} \tag{11}$$

where the activation function, $f(x)$, is defined in equations (2) and (4). The first of these equations is for weights between the middle and output layers, Z_{jk} and Y_k. The size of the updates depends on the difference between the network's node activation and the desired output, the derivative of the activation function (which tends to suppress weight changes when a node saturated one way or the other near 0 or 1), and the size of the output from the middle layer, b_j. The second of these equations, (11), is for weights between the input and middle layers, W_{ij} and X_j. It represents a recursive computation to propagate the error backward through the network.

Three sets of weights corresponding to the three stages in Figure 19 are determined for each station. Separate data sets were used for training and testing (2/3 of the available data were randomly selected for training and the other 1/3 was used for testing). The number of input patterns for each 3-component IRIS station was typically about 300, and these patterns were presented to the neural network approximately 1000 times (although this varied from station to station). This represents data from about one week of continuous operation. The learning rate was typically set to about 0.1 and the momentum was set to about 0.5. The total training time for each station (all three stages) is of the order of 45 minutes on a Sun Sparc-2 workstation.

III. Operational Test and Evaluation

This section gives the test and evaluation results for neural networks trained for six 3-component IRIS/IDA stations in the CIS, and a 4-element micro-array in Kislovodsk near the Caucasus Mountains.

3.1 Three-Component Data

Neural networks were developed and trained for six 3-component, broadband IRIS/IDA stations in the CIS: Ala-Archa (AAK) in Kyrgyzstan, Arti (ARU) in the southern Ural Mountains, Garm (GAR) in Tadjikistan, Kislovodsk (KIV) near the Caucasus Mountains, Obninsk (OBN) near Moscow, and Talaya (TLY) near Lake Baikal [*Given*, 1990; *Given*, 1991; *Given and Fels*, 1993]. The station locations are shown in Figure 21. Data recorded during the period, 11-17 July, 1991, that were processed by *IMS* and analyzed by F. Ryall at the Center for Seismic Studies (CSS) were used to train the neural networks. Since several of these stations had few analyst-reviewed data, a set of average weights was also derived by training a neural network with data from all stations. The input data for the neural networks are plotted in Figures 2-16.

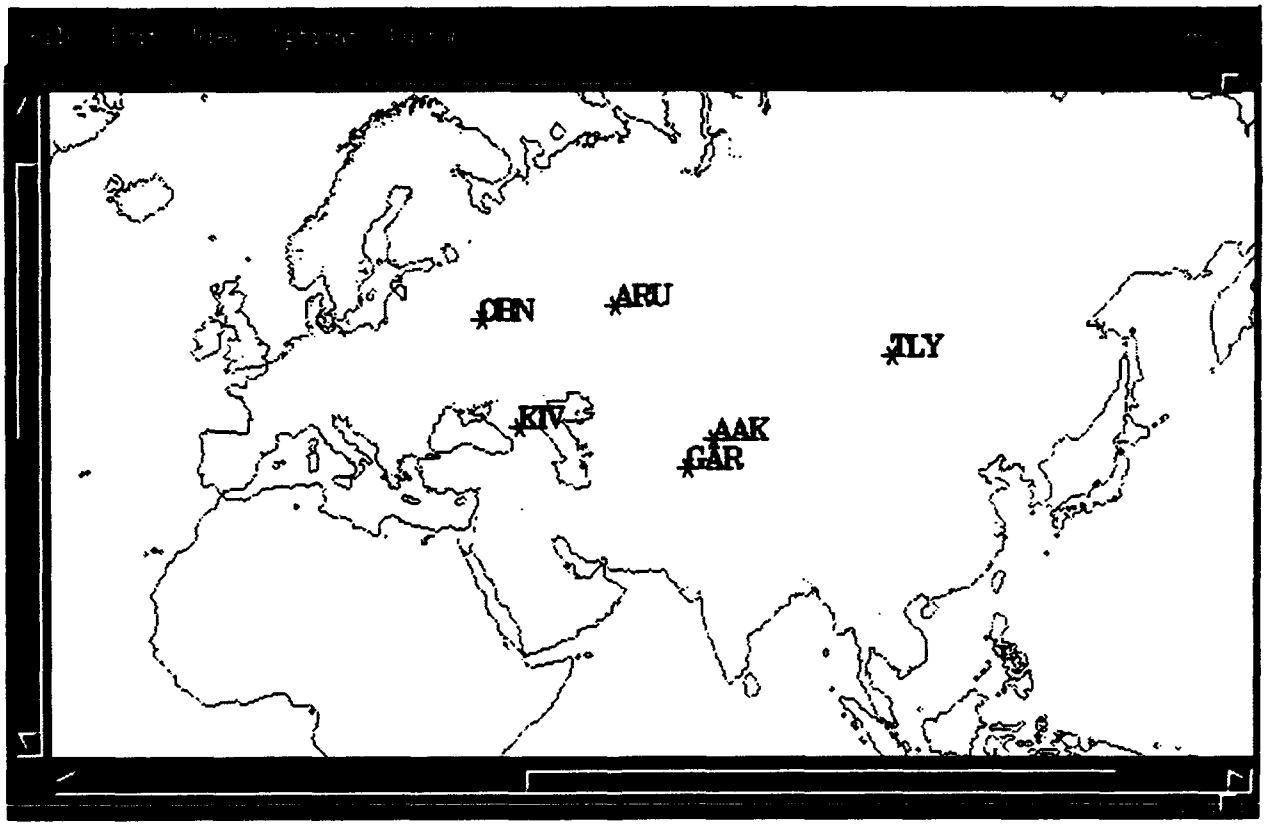


Figure 21. This map shows the locations of six 3-component IRIS/IDA stations used in this study.

The testing results are shown in the *confusion matrices* in Tables 1-7 (next page). These tables compare the initial wave type determined by the analyst (rows) to that determined by the neural network (columns). The diagonal elements are the numbers of arrivals for which the neural network produced the same identification as the analyst (i.e., correct identification). Similarly, the off-diagonal elements are the numbers of arrivals for which the neural network produced a different identification than the analyst (i.e., incorrect identification). The average identification accuracy for each station is included in the table heading. This is equal to the sum of the diagonal elements divided by the sum of all elements. Only one station has an identification accuracy less than 90% (the accuracy at GAR is 89.6%). The average identification accuracy of the neural network that was trained using data from all IRIS/IDA stations (i.e., average weights) is 77%.

The adaptability of our neural networks to data recorded in various geologic environments was tested using the approach that we developed for our previous two-class (*P* or *S*) neural network [*Patnaik and Sereno*, 1991]. That is, we computed the identification accuracy at each station produced by neural networks that were trained with data from other stations. The results are shown in Table 8. The diagonal elements are the results of testing and training with data from the same station (i.e., these are the same as the percentages given in the headings of Tables 1-7). The off-diagonal elements are the results of cross-testing (i.e., adaptability testing). A trained neural network generally shows only about 50% identification accuracy if applied to data from a new station, without retraining. The identification accuracy is generally >90% if testing and training is performed with data from the same station. The cross-testing was more successful for our 2-class (*P* or *S*) neural networks. For these, the off-diagonal terms were about 80%. The difference appears to be related to the strong station-dependence in the characteristics of the noise.

Table 8: ADAPTABILITY TESTING

Train \ Test	ALL	AAK	ARU	GAR	KIV	OBN	TLY
ALL	**77.2**	82.5	60.9	80.1	75.8	95.7	73.9
AAK	45.3	**95.4**	47.2	55.9	19.9	10.2	47.2
ARU	70.4	29.4	**95.4**	43.8	67.1	95.3	60.8
GAR	65.9	62.3	51.7	**89.6**	39.1	68.1	56.3
KIV	63.3	21.3	76.2	37.4	**94.3**	94.9	48.2
OBN	64.3	21.0	85.3	34.2	71.3	**98.4**	50.3
TLY	63.0	57.0	57.8	57.3	51.7	75.6	**98.5**

The first operational test and evaluation of our new 4-class neural network was conducted by applying **ESAL** to the one-week IRIS data set. **ESAL** was applied twice; once using the trained neural networks to identify initial wave type, and once using the rule-based system (Figure 1). Analyst-reviewed event solutions from the *IMS* array stations (NORESS, ARCESS, FINESA, and

Table 1: AAK (95.4%)

Analyst \ Neural Net	T	P	S	N
T	44	1	1	7
P	1	104	1	0
S	0	0	115	0
N	5	1	1	114

Table 2: ARU (95.4%)

Analyst \ Neural Net	T	P	S	N
T	69	0	0	24
P	1	2	0	1
S	0	0	6	8
N	15	0	1	954

Table 3: GAR (89.6%)

Analyst \ Neural Net	T	P	S	N
T	58	0	1	20
P	1	232	1	10
S	0	2	249	7
N	6	30	27	364

Table 4: KIV (94.3%)

Analyst \ Neural Net	T	P	S	N
T	24	0	0	0
P	0	25	0	0
S	0	0	31	0
N	4	8	3	166

Table 5: OBN (98.4%)

Analyst \ Neural Net	T	P	S	N
T	28	0	0	0
P	0	11	0	0
S	0	0	5	1
N	4	2	5	694

Table 6: TLY (98.5%)

Analyst \ Neural Net	T	P	S	N
T	41	0	0	2
P	1	16	0	0
S	0	0	32	0
N	0	0	0	107

Table 7: ALL IRIS STATIONS (77.2%)

Analyst \ Neural Net	T	P	S	N
T	260	6	13	41
P	6	369	7	25
S	5	3	405	44
N	271	101	319	1820

GERESS) were used as starting solutions for the automated interpretation of the IRIS data (i.e., the IRIS data provided an incremental addition to an existing bulletin). The results are given in Table 9 for events recorded by at least one of the 3-component IRIS stations. The first column is for *IMS* Version 2 which uses the rule-based system to identify initial wave type, and the second column is for *IMS* Version 3 which uses the trained neural networks.

Table 9: PERFORMACE SUMMARYFOR 1-WEEK IMS+IRIS DATA SET

	IMS Version 2	*IMS* Version 3
Analyst Events	265	265
ESAL Events	1069	381
Moved <50 km	118	129
Moved >50 km	92	84
Added Events	55	52
Rejected Events	859	168

The first row in Table 9 is the number of events in the analyst's bulletin, and the second row is the number of events formed by **ESAL** during the one-week test period. The third row is the number of events whose location solution determined by **ESAL** differs from the analyst's location solution by <50km. Similarly, the fourth row is the number of events for which **ESAL**'s location solution differs from the analyst's location solution by >50 km. The fifth row is the number of events that are in the analyst's bulletin that are not in **ESAL**'s bulletin (i.e., events added by the analyst and missed by **ESAL**). The sixth row is the number of events in **ESAL**'s bulletin that are not in the analyst's bulletin (i.e. events rejected by the analyst as false-alarms). The most obvious difference between the two versions is the number of false-alarms. Application of the neural network (*IMS* Version 3) reduced these by >80%.

A second operational test was conducted using a larger data set that did not include the data used to train the neural networks. For this test, *IMS* Version 2 and *IMS* Version 3 were applied to different 3-week data sets as part of a cyclical test conducted at CSS. The purpose of this test was to demonstrate improvement in the performance of *IMS* for 3-component data. *IMS* Version 2 was used for the first 3-week data set recorded between July 6 and July 26, 1991. The analyst-reviewed data from this test were used to develop and train the 4-class neural networks. We only used data from one of the three weeks for training because the analyst only identified noise detections for that week. *IMS* Version 3 was applied to the second 3-week data set recorded between July 27 and August 16, 1991. This was the first operational test of the neural networks that used data that were not included in the training. Table 10 compares the results of the two tests. As was the case for the one-week data set, the false-alarms were reduced significantly by *IMS* Version 3 (by about 60% for this data set). We attribute this success to the ability of the neural network to perform accurate noise screening.

Table 10: PERFORMANCE SUMMARY FOR CYCLICAL OPERATIONS AT CSS

	IMS Version 2 Cycle 1	*IMS* Version 3 Cycle 2
Analyst Events	634	723
ESAL Events	2721	1481
Moved <50 km	257	265
Moved >50 km	231	274
Added Events	146	184
Rejected Events	2233	942

3.2 Micro-array Data

Micro-arrays were originally suggested by *Kværna and Ringdal* [1990] as a less-costly alternative to the larger NORESS-type arrays that still provided slowness and azimuth estimates accurate enough for reliably identifying initial wave type and grouping arrivals that belong to the same event. In particular, they found that *f-k* analysis of NORESS A-ring data (the radius is about 150 m) provided adequate resolution to achieve an accuracy of 95% for automated identification of *P* and *S* phases, and azimuth uncertainties of about 30° for regional *P* and *S* phases. Based on these results, Scripps Institution of Oceanography (SIO), the Center for Seismic Studies (CSS), and the Experimental Methodical Expedition (EME) in Obninsk collaborated on an experiment to establish and operate a 4-element micro-array in Kislovodsk [*Berger et al.*, 1992]. The center element is a 3-component seismometer, and the other three elements are vertical-component seismometers. The arms of the array are 150 m in length, and they are 120° apart (Figure 22).

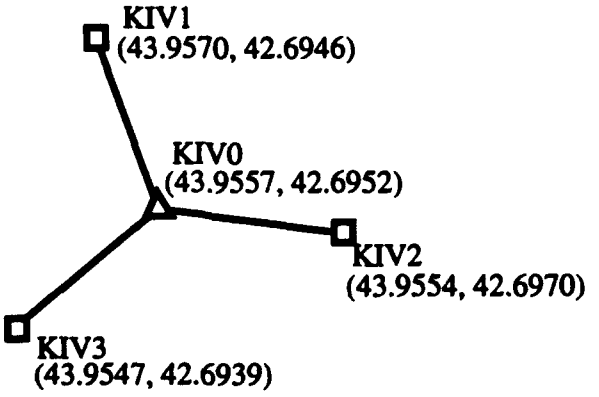


Figure 22. The geometry of the 4-element micro-array in Kislovodsk is plotted.

We analyzed a two-week data set from the micro-array in Kislovodsk (KIV0) recorded in October, 1992. While this data set is too small for definitive conclusions, our preliminary results suggest that the slowness estimates from this array are not accurate enough for reliable identification of initial wave type (see the last column in Figure 17). In contrast, our neural network achieves >90% identification accuracy for all wave-types (Table 11).

Table 11: KIV0 (93.5%)

Analyst \ Neural Net	T	P	S	N
T	**97**	1	1	4
P	2	**93**	0	4
S	2	3	**83**	5
N	7	4	5	**232**

The neural network includes all the input attributes for a 3-component station plus the *f-k* slowness estimate, but the results are comparable to those for 3-component stations (i.e., slowness from this array does not contribute much to the initial wave-type identification). The main advantage of the micro-array over a 3-component station is that the array provides azimuth estimates for regional *S* waves to aid the grouping of arrivals from the same event. However, the KIV0 azimuth uncertainties are much larger than they are for NORESS-type arrays. For example, for typical regional signals the azimuth uncertainty is about 7° for NORESS-type arrays and about 35° for KIV0. While this uncertainty is too large to be useful for location, it may be accurate enough to reduce the number of false events formed by **ESAL**. A much larger analyst-verified data set is needed to investigate this conjecture. Such a data set is currently being compiled at CSS for KIV0 and two 9-element mini-arrays (Apatity and Spitsbergen). These data will be used to quantify the dependence of the false-alarm rate on the array geometry.

IV. Conclusions

We developed and implemented a neural network technique in *IMS* for automated identification of the initial wave type of seismic phases (Teleseism, Regional *P*, Regional *S*, and Noise) recorded by 3-component stations and arrays. This is an extension of the 2-class neural network (*P* or *S*) that we developed for 3-component stations [*Patnaik and Sereno*, 1991]. The identification accuracies of the 4-class neural networks are >90% for most of the stations that we tested. The average results for data from six 3-component IRIS/IDA stations in the CIS are:

$$T = 82.5\% \quad P = 95.8\% \quad S = 96.1\% \quad N = 95.5\%$$

when separate weights are derived for each station. The key advantages of the neural network approach listed by *Patnaik and Sereno* [1991] for the 2-class neural network also apply to the new 4-class neural network. These include:

- It is easier to develop than rules because initial wave-type identification is based on high-dimensional multivariate input data (e.g., 15 -16 input attributes).
- It is easily extended to include new features, which may be difficult in a conventional rule-based system.
- It performs better than linear multivariate methods.
- It incorporates station-specific characteristics.
- It is easily adapted to data from new stations (enough data for retraining can be accumulated in about 1-2 weeks of continuous station operation, and training generally takes less than one hour on a Sun Sparc-2 workstation).

In addition, the new neural network has been extended for application to arrays by including *f-k* slowness to the input attributes. This was motivated by the integration of data from mini- and micro-arrays into the *IMS* whose *f-k* resolution is not sufficient to reliably identify initial wave type.

The 4-class neural network was implemented in *IMS* and tested under operational conditions using data from the 3-component IRIS/IDA stations in the CIS. The key result was a reduction in the number of false-alarms produced by the automated system by about 60%, and this was accomplished without a significant increase in the number of missed events. We attribute this success to the ability of the neural network to perform accurate noise-screening for data recorded by 3-component stations.

Acknowledgments

We gratefully acknowledge the help of several of our colleagues at SAIC and Fred Dashiell at Inference Corporation. We especially thank Rick Jenkins and Fred Dashiell for their work in integrating the neural network software module into **ESAL**. Rick Jenkins also performed most of the operational testing, and he provided valuable contributions to the evaluation. We are also grateful to Tim Ash for his help in training the neural networks, and for his comments and suggestions regarding the neural network architecture. We thank Jeff Given who developed the spectral parameterization of horizontal-to-vertical power ratio, and we thank Christine Ferraro for her assistance in the final production of this report.

References

Bache, T., J. Anderson, D. Baumgardt, S. Bratt, W. Farrell, R. Fung, J. Given, A. Henson, C. Kobryn, H. Swanger, J. Wang, Intelligent Array System, Final Technical Report, SAIC-90/1431, San Diego, California, 170 pp., 1990*a*.

Bache, T., S. Bratt, J. Given, T. Schroeder, H. Swanger, J. Wang, The Intelligent Monitoring System Version 2, Quarterly Technical Report #7, SAIC-91/1137, San Diego, California, 93 pp., 1991.

Bache, T., S. Bratt, H. Swanger, G. Beall, F. Dashiell, Knowledge-based interpretation of seismic data in the Intelligent Monitoring System, *Bull. Seismol. Soc. Am.*, *83*, 1507-1526, 1993.

Bache, T., S. Bratt, J. Wang, R. Fung, C. Kobryn, and J. Given, The Intelligent Monitoring System, *Bull. Seismol. Soc. Am.*, *80, Part B*, 1833-1851, 1990*b*.

Berger, J., D. Chavez, J. Given, and D. Williams, Kislovodsk mini-array experiment, October 7 through October 20, 1992, Final Report, Institute of Geophysics and Planetary Physics, La Jolla, California, 69 pp., 1992.

Bratt, S., G. Beall, H. Swanger, F. Dashiell, and T. Bache, A knowledge-based system for automatic interpretation of seismic data to associate signals and locate events, Quarterly Technical Report #8, SAIC-91/1281, San Diego, California, 73 pp., 1991.

Given, H., Variations in broadband noise at IRIS/IDA stations in the USSR with implications for event detection, *Bull. Seismol. Soc. Am.*, *80*, 2072-2088, 1990.

Given, H., Overview of Eurasian seismic stations and networks, in the Proceedings of the 13th Annual PL/DARPA Seismic Research Symposium, 8-10 October, PL-TR-91-2208, 3-12, 1991.

Given, H. and J. Fels, Site characteristics and ambient ground noise at IRIS/IDA stations AAK (Ala-Archa, Kyrgyzstan) and TLY (Talaya, Russia), *Bull. Seismol. Soc. Am.*, *83*, 945-952, 1993.

Jurkevics, A., Polarization analysis of 3-component array data, *Bull. Seismol. Soc. Am.*, *78*, 1725-1743, 1988.

Kværna, T. and F. Ringdal, Real-time processing using a hybrid 3-component/small array station, Semiannual Tech. Summary, NORSAR Sci. Rep. No. 1-90/91, 96-109, 1990.

Mykkeltveit, S., A. Dahle, J. Fyen, T. Kværna, P. Larsen, R. Paulsen, F. Ringdal, and I. Kuzmin, Extensions of the northern Europe regional array network - New small-aperture arrays in Apatity, Russia, and on the Arctic island of Spitsbergen, NORSAR Scientific Report No. 1-92/93, Kjeller, Norway, 58-71, 1992.

Patnaik, G. and T. Sereno, Neural computing for seismic phase identification, PL-TR-92-2110(II), 42 pp., 1991.

Patnaik, G., T. Sereno, and R. Jenkins, Test and evaluation of neural network applications for seismic signal discrimination, PL-TR-92-2218 (II), 1992.

Rumelhart, D., J. McClelland, and the PDP Research Group, *Parallel Distributed Processing, Volume 1: Foundations*, The MIT Press, Cambridge, Massachusetts, 1986.

Suteau-Henson, A., Three-component analysis of regional phases at NORESS and ARCESS: Polarization and phase identification, *Bull. Seismol. Soc. Am.*, *81*, 2419-2440, 1991.

Swanger, H., J. Anderson, T. Sereno, J. Given, and D. Williams, The Intelligent Monitoring System extensions to the Center Version 3 database (Rev. 1), Technical Report, SAIC-93/1123, 106 pp., 1993.

DISTRIBUTION LIST

RECIPIENT	NUMBER OF COPIES

DEPARTMENT OF DEFENSE

ARPA/NMRO — 3
ATTN: Dr. R. Alewine, Dr. S. Bratt, and Dr. A. Ryall, Jr.
3701 North Fairfax Drive
Arlington, VA 22203-1714

ARPA, OASB/Library — 1
3701 North Fairfax Drive
Arlington, VA 22203-1714

Defense Technical Information Center — 2
Cameron Station
Alexandria, VA 22314

DEPARTMENT OF THE AIR FORCE

AFTAC/TT — 3
ATTN: Dr. L. Himes, Dr. F. Pilotte, and Dr. D. Russell
130 South Highway A1A
Patrick AFB, FL 32925-3002

AFTAC/TT, Center for Seismic Studies — 1
ATTN: Dr. R. Blandford
1300 North 17th Street, Suite 1450
Arlington, VA 22209-2308

Phillips Laboratory/GPEH — 1
ATTN: Mr. J. Lewkowicz
29 Randolph Road
Hanscom AFB, MA 01731-3010

DEPARTMENT OF ENERGY

Department of Energy — 1
ATTN: Dr. M. Denny
Office of Arms Control
Washington, D.C. 20585

Lawrence Livermore National Laboratory 4
ATTN: Dr. J. Hannon, Dr. K. Nakanishi, Dr. H. Patton,
and Dr. D. Springer
University of California
P.O. Box 808
Livermore, CA 94550

Los Alamos National Laboratory 1
ATTN: Dr. S. Taylor
P.O. Box 1663, Mail Stop C335
Los Alamos, NM 87545

Sandia National Laboratory 2
ATTN: Dr. E. Chael and Dr. M. Sharp
Division 9241
Albuquerque, NM 87185

OTHER GOVERNMENT AGENCIES

Central Intelligence Agency 1
ATTN: Dr. L. Turnbull
CIA-OSWR/NED
Washington, D.C. 20505

U.S. Geological Survey 1
ATTN: Dr. A. McGarr
Mail Stop 977
Menlo Park, CA 94025

U.S. Geological Survey 1
ATTN: Dr. W. Leith
Mail Stop 928
Reston, VA 22092

U.S. Geological Survey 1
ATTN: Dr. R. Masse
Denver Federal Building
Box 25046, Mail Stop 967
Denver, CO 80225

UNIVERSITIES

Boston College 1
ATTN: Dr. A. Kafka
Department of Geology and Geophysics
Chestnut Hill, MA 02167

California Institute of Technology 1
ATTN: Dr. D. Helmberger
Seismological Laboratory
Pasadena, CA 91125

Columbia University 2
ATTN: Dr. P. Richards and Dr. L. Sykes
Lamont-Doherty Geolo cal Observatory
Palisades, NY 10964

Cornell University 1
ATTN: Dr. M. Barazangi
Institute for the Study of the Continent
Ithaca, NY 14853

IRIS, Inc. 2
ATTN: Dr. D. Simpson and Dr. G. van der Vink
1616 North Fort Myer Drive, Suite 1050
Arlington, VA 22209

Massachusetts Institute of Technology 1
ATTN: Dr. T. Jordan
Department of Earth, Atmospheric and Planetary Sciences
Cambridge, MA 02139

Massachusetts Institute of Technology 1
ATTN: Dr. N. Toksoz
Earth Resources Laboratory
42 Carleton Street
Cambridge, MA 02142

MIT Lincoln Laboratory, M-200B 1
ATTN: Dr. R. Lacoss
P.O. Box 73
Lexington, MA 02173-0073

San Diego State University 1
ATTN: Dr. S. Day
Department of Geological Sciences
San Diego, CA 92182

Southern Methodist University 2
ATTN: Dr. E. Herrin and Dr. B. Stump
Institute for the Study of Earth and Man
Geophysical Laboratory
Dallas, TX 75275

Southern Methodist University 1
ATTN: Dr. Gary McCartor
Department of Physics
Dallas, TX 75275

State University of New York at Binghamton 2
ATTN: Dr. J. Barker and Dr. F. Wu
Department of Geological Sciences
Vestal, NY 13901

St. Louis University 2
ATTN: Dr. R. Herrmann and Dr. B. Mitchell
Department of Earth and Atmospheric Sciences
St. Louis, MO 63156

The Pennsylvania State University 2
ATTN: Dr. S. Alexander and Dr. C. Langston
Geosciences Department
403 Deike Building
University Park, PA 16802

University of Arizona 1
ATTN: Dr. T. Wallace
Department of Geosciences, Building #77
Tucson, AZ 85721

University of California, Berkeley 2
ATTN: Dr. L. Johnson and Dr. T. McEvilly
Seismographic Station
Berkeley, CA 94720

University of California, Davis 1
ATTN: Dr. R. Shumway
Division of Statistics
Davis, CA 95616

University of California, San Diego 5
ATTN: Dr. J. Berger, Dr. L. Burdick, Dr. H. Given, Dr. B. Minster,
and Dr. J. Orcutt
Scripps Institute of Oceanography, A-025
La Jolla, CA 92093

University of California, Santa Cruz 1
ATTN: Dr. T. Lay
Institute of Tectonics
Earth Science Board
Santa Cruz, CA 95064

University of Colorado 2
ATTN: Dr. C. Archambeau and Dr. D. Harvey
CIRES
Boulder, CO 80309

University of Connecticut 1
ATTN: V. Cormier
Department of Geology and Geophysics
U-45, Room 207
Storrs, CT 06268

University of Southern California 1
ATTN: Dr. K. Aki
Center for Earth Sciences
University Park
Los Angeles, CA 90089-0741

University of Wisconsin-Madison 1
ATTN: Dr. C. Thurber
Department of Geology and Geophysics
1215 West Dayton Street
Madison, WS 53706

DEPARTMENT OF DEFENSE CONTRACTORS

Center for Seismic Studies 3
ATTN: Dr. R. Bowman, Dr. J. Carter, and Dr. R. Gustafson
1300 North 17th Street, Suite 1450
Arlington, VA 22209

ENSCO, Inc. 2
ATTN: Dr. D. Baumgardt and Dr. Z. Der
5400 Port Royal Road
Springfield, VA 22151-2388

ENSCO, Inc. 2
ATTN: Dr. R. Kemerait and Dr. D. Taylor
445 Pineda Court
Melbourne, FL 32940-7508

Mission Research Corporation 1
ATTN: Dr. M. Fisk
735 State Street
PO Drawer 719
Santa Barbara, CA 93102-0719

Radix Systems, Inc. 1
ATTN: Dr. J. Pulli
201 Perry Parkway
Gaithersburg, MD 20877

Science Horizons 1
ATTN: Dr. T. Cherry
710 Encinitas Blvd., Suite 200
Encinitas, CA 92024

S-CUBED, 2
A Division of Maxwell Laboratory
ATTN: Dr. T. Bennett and Mr. J. Murphy
11800 Sunrise Valley Drive, Suite 1212
Reston, VA 22091

S-CUBED, 2
A Division of Maxwell Laboratory
ATTN: Dr. K. McLaughlin and Dr. J. Stevens
P.O. Box 1620
La Jolla, CA 92038-1620

SRI International 2
ATTN: Dr. A. Florence and Dr. S. Miller
333 Ravenswood Avenue, Box AF116
Menlo Park, CA 94025-3493

Teledyne Geotech 1
ATTN: Mr. W. Rivers
314 Montgomery Street
Alexandria, VA 22314-1581

TASC, Inc. 1
ATTN: Dr. R. Comer
55 Walkers Brook Drive
Reading, MA 01867

NON-US RECIPIENTS

Blacknest Seismological Center 1
ATTN: Dr. P. Marshall
UK Ministry of Defense
Blacknest, Brimpton
Reading FG7-FRS, UNITED KINGDOM

Institute for Geophysik 1
ATTN: Dr. H.-P. Harjes
Ruhr University/Bochum
P.O. Box 102148
4630 Bochum 1, GERMANY

NTNF/NORSAR 2
ATTN: Dr. S. Mykkeltveit and Dr. F. Ringdal
P.O. Box 51
N-2007 Kjeller, NORWAY

Societe Radiomana 1
ATTN: Dr. B. Massinon
27 Rue Claude Bernard
75005 Paris, FRANCE

University of Cambridge 1
ATTN: Dr. K. Priestley
Bullard Labs, Department of Earth Sciences
Madingley Rise, Madingley Road
Cambridge CB3, OEZ, ENGLAND

University of Toronto 1
ATTN: Dr. K.-Y. Chun
Geophysics Division
Physics Department
Ontario, CANADA